

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 26, 2005, 11:54:19 ; Search time 73.3055 Seconds
(without alignments)
153.004 Million cell updates/sec

Title: US-10-624-218-1

Perfect score: 29

Sequence: 1 RAAFLEXENTALTREVAEAEKEVRCENI 29

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 2105692 seqs, 386760381 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	29	100.0	29	AD134276	Adi34276 Zipper pe
2	19	65.5	261	ADD47845	Add47845 Rat Prote
3	19	65.5	303	ADD47847	Add47847 Human Pro
4	19	65.5	303	ADJ68701	Adj68701 Human hea
5	19	65.5	303	AD182509	Adi82509 Human mod
6	19	65.5	303	ABM82367	Abm82367 Tumour-as
7	14	48.3	29	AD134277	Adi34277 Zipper pe
8	12	41.4	43	AD134279	Adi34279 Zipper pe
9	9	31.0	325	AD56596	Ades6596 Human Pro
10	9	31.0	325	AD46912	Add46912 Rat Prote
11	9	31.0	325	AD56594	Ades6594 Rat Prote
12	9	31.0	325	AD46916	Add46916 Rat Prote
13	9	31.0	325	AD46914	Add46914 Human Pro
14	9	31.0	325	AD46918	Add46918 Human Pro
15	9	31.0	325	AD56590	Ades6590 Rat Prote
16	9	31.0	325	AD56592	Ades6592 Human Pro
17	9	31.0	325	AD182507	Adi82507 Human mod
18	9	31.0	325	AD019734	Ado19734 Human PRO
19	9	31.0	325	ADP54859	Adp54859 Human PRO
20	8	27.6	202	AA827986	Aab27986 Human sec
21	8	27.6	203	AA827985	Aab27985 Human sec
22	8	27.6	334	AA803211	Aar03211 Amino aci
23	8	27.6	413	AAW46822	Aaw46822 Amino aci
24	8	27.6	534	AAW46823	Aaw46823 Amino aci
25	8	27.6	534	AAW98948	Aaw98948 Streptoco

26	8	27.6	539	8	ADL90009	Adl90009 Gluconoba
27	8	27.6	567	8	ABU43677	Abu43677 Protein e
28	8	27.6	1068	7	ADE08494	Ado08494 Novel pro
29	8	27.6	1068	8	ADS10806	Adsl0806 Human the
30	8	27.6	1077	4	AAW78896	Aam78896 Human pro
31	8	27.6	1106	4	AAW79880	Aam79880 Human pro
32	8	27.6	1106	4	AAW79880	Aam79880 Human pro
33	7	24.1	163	7	ADB69976	Adb69976 C. neofo
34	7	24.1	284	8	ADN21004	Adn21004 Bacterial
35	7	24.1	287	5	ABB48469	Abb48469 Listeria
36	7	24.1	357	6	ABU26605	Abu26605 Protein e
37	7	24.1	388	5	ABB50100	Abb50100 Listeria
38	7	24.1	431	7	ADH87467	Adh87467 Enterococ
39	7	24.1	455	4	ASG07271	Asg07271 Novel hum
40	7	24.1	1084	8	AD110225	Adi10225 Interrupt
41	6	20.7	1873	4	ABG14982	Abg14982 Novel hum
42	6	20.7	8	6	ABR58097	AbR58097 HIV rever
43	6	20.7	8	6	ABR58098	AbR58098 HIV rever
44	6	20.7	9	6	ABR58099	AbR58099 HIV rever
45	6	20.7	9	6	ABM66287	Abm66287 Propionib
					ABR58108	ABR58108 HIV rever

ALIGNMENTS

RESULT 1
AD134276
ID AD134276 standard; protein; 29 AA.
XX
AC AD134276;
XX
DT 15-APR-2004 (first entry)
XX
DE Zipper peptide #1 for cross linking adenoviral ligands.
XX
KW adenovirus vector; Cytostatic; HER2/neu; CD40; tumor necrosis factor;
KW TNF.
XX
OS Unidentified.
XX
PN WO2004009133-A1.
XX
PD 29-JAN-2004.
XX
PF 22-JUL-2003; 2003WO-US022852.
XX
PR 22-JUL-2002; 2002US-037951P.
XX
PA (VECT-) VECTORLOGICS INC.
XX
PI Korokhov N, Mikhveva G;
XX
DR WPI; 2004-132871/13.
XX
PT Novel recombinant adenovirus having fiber protein modified by insertion
PT of first zipper peptide that can crosslink to second zipper peptide-
PT targeting ligand fusion protein, and binding between zipper peptides
PT targets vector to cell.
XX
PS Claim 2; SEQ ID NO 1; 54pp; English.

The present invention relates to a targeted recombinant adenovirus vector. The invention is useful for expressing a heterologous protein chosen from a tumor associated antigen, HER2/neu and carcinoembryonic antigen, in a target cell e.g., CD40 + cell such as dendritic cells, is useful in a method of gene transfer to CD40 + cells, where the targeted adenovirus vector mediates transfer of the gene encoding heterologous protein to the cell such as the dendritic cell. The vector is useful in gene therapy techniques for treatment of tumors. Multivalent interaction or trimeric CD40L with CD40 receptors causes CD40 ligation, which then results in enhanced survival of these cells and secretion of cytokines such as interleukin (IL)-1, IL-6, IL-8, IL-10, IL-12, tumor necrosis factor (TNF)- α agr γ , macrophage inflammatory protein (MIP)-1 α and enzymes

CC such as matrix metalloproteinase. CD40-CD40L interaction also enhances
CC monocyte tumoricidal activity. The present sequence represents a zipper
CC peptide to be used for cross-linking targeting ligands with adenoviral
CC variants.
XX
XX
SQ Sequence 29 AA;
Query Match 100.0%; Score 29; DB 8; Length 29;
Best Local Similarity 100.0%; Pred. No. 2.2e-21;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 RAAFLKENTALTREVAELKEVGRCENI 29
Db 1 RAAFLKENTALTREVAELKEVGRCENI 29
RESULT 2
ADD47845
ID ADD47845 standard; protein; 261 AA.
XX
AC ADD47845;
XX
XX 02-DEC-2004 (revised)
DT 29-JAN-2004 (first entry)
XX
XX Rat Protein AAB20032, SEQ ID NO 13541.
DE
XX Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
XX
XX Rattus norvegicus.
OS Unidentified.
XX
XX WO2003016475-A2.
PN
XX 27-FEB-2003.
PD
XX
XX 14-AUG-2002; 2002WO-US025765.
PF
XX
XX 14-AUG-2001; 2001US-0312147P.
PR
XX 01-NOV-2001; 2001US-0346382P.
PR
XX 26-NOV-2001; 2001US-0333347P.
XX
XX (GEO) GEN HOSPITAL CORP.
PA (FARB) BAYER AG.
XX
XX Woolf C, D'urso D, Befort K, Costigan M;
PI
XX
XX WPI; 2003-268312/26.
DR GENBANK; AAB20032.
XX
XX New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
XX
XX Example 1; Page; 1017pp; English.
PS
XX
XX The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating

CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a rat protein (described in Table 3
CC of the specification) which is differentially expressed during pain.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 261 AA;
Query Match 65.5%; Score 19; DB 7; Length 261;
Best Local Similarity 100.0%; Pred. No. 1e-10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RAAFLKENTALTREVAEL 19
Db 222 RAAFLKENTALTREVAEL 240

RESULT 3
ADD47847
ID ADD47847 standard; protein; 303 AA.
XX
AC ADD47847;
XX
XX 02-DEC-2004 (revised)
DT 29-JAN-2004 (first entry)
XX
XX Human Protein NP_003207, SEQ ID NO 13543.
DE
XX Human; pain; neuronal tissue; gene therapy;
KW spinal segmental nerve injury; chronic constriction injury; CCI;
KW spared nerve injury; SNI; Chung.
XX
XX Homo sapiens.
OS Unidentified.
XX
XX WO2003016475-A2.
PN
XX
XX 27-FEB-2003.
PD
XX
XX 14-AUG-2002; 2002WO-US025765.
PF
XX
XX 14-AUG-2001; 2001US-0312147P.
PR
XX 01-NOV-2001; 2001US-0346382P.
PR
XX 26-NOV-2001; 2001US-0333347P.
XX
XX (GEO) GEN HOSPITAL CORP.
PA (FARB) BAYER AG.
XX
XX Woolf C, D'urso D, Befort K, Costigan M;
PI
XX
XX WPI; 2003-268312/26.
DR GENBANK; NP_003207.
XX
XX New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
XX
XX Example 1; Page; 1017pp; English.

CC The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates

CC the expression of a polynucleotide sequence which is differentially
 CC expressed in an animal subjected to pain, a method for identifying a
 CC compound that regulates the activity of one or more of the
 CC polynucleotides, a method for producing a pharmaceutical composition, a
 CC method for identifying a compound or small molecule that regulates the
 CC activity in an animal of one or more of the polypeptides given in the
 CC specification, a method for identifying a compound useful in treating
 CC pain and a pharmaceutical composition comprising the one or more
 CC polypeptides or their antibodies. The polynucleotide or the compound that
 CC modulates its activity is useful for preparing a medicament for treating
 CC pain (e.g. spinal segmental nerve injury (SNI)) in an animal (e.g. gene
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
 CC therapy). The sequence presented is a human protein (described in Table 3
 CC of the specification) which is differentially expressed during pain.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic form directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 303 AA;

Query Match 65.5%; Score 19; DB 7; Length 303;
 Best Local Similarity 100.0%; Pred. No. 1.2e-10;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RAAFLKENTALTREVAEL 19
 |||||
 Db 264 RAAFLKENTALTREVAEL 282

RESULT 4

ADJ68701
 ID ADJ68701 standard; protein; 303 AA.

XX AC ADJ68701;

XX DT 06-MAY-2004 (first entry)

XX DE Human heat mitochondrial protein as a therapeutic target SeqID507.

XX KW mitochondrial; human; screening assay; diabetes mellitus;

XX KW Huntington's disease; osteoarthritis;

XX KW Leber's hereditary optic neuropathy; LHON;

XX KW mitochondrial encephalopathy lactic acidosis and stroke; MELAS;

XX KW myoclonic epilepsy ragged red fibre syndrome; MERRF; cancer;

XX KW neuroprotective; nootropic; antidiabetic; anticonvulsant; antiarthritic;

XX KW osteopathic; ophthalmological; cyostatic.

XX OS Homo sapiens.

XX PN WO2003087768-A2.

XX PD 23-OCT-2003.

XX PF 04-APR-2003; 2003WO-US010870.

XX PR 12-APR-2002; 2002US-0372843P.

XX PR 17-JUN-2002; 2002US-038987P.

XX PR 20-SEP-2002; 2002US-0412418P.

XX PA (MITO-) MITOKOR.

XX PA (BUCK-) BUCK INST AGE RES.

XX PI Ghosh SS, Fahy ED, Zhang B, Gibson BW, Taylor SW, Glenn GM;

XX PI Warnock DE;

XX DR WPI; 2003-845369/78.

XX PT Identifying a mitochondrial target for drug screening assays and for
 PT treating diseases associated with altered mitochondrial function,
 PT comprises detecting a modified polypeptide in a sample and correlating
 PT with the disease.

XX PS Claim 1; SEQ ID NO 507; 180pp; English.

XX CC This invention relates to novel mitochondrial targets that can be used
 CC for therapeutic intervention in treating a disease associated with
 CC altered mitochondrial function. Specifically, it refers to a method for
 CC identifying proteins of the human heart mitochondrial proteome that are
 CC useful for drug screening assays, as well as therapeutic targets. The
 CC present invention describes a method for identifying such proteins that
 CC can be used in the treatment of various diseases associated with altered
 CC mitochondrial function including diabetes mellitus, Huntington's disease,
 CC osteoarthritis, Leber's hereditary optic neuropathy (LHON), mitochondrial
 CC encephalopathy lactic acidosis and stroke (MELAS), myoclonic epilepsy
 CC ragged red fibre syndrome (MERRF) or cancer. Accordingly, these
 CC compositions have neuroprotective, nootropic, antidiabetic,
 CC anticonvulsant, antiarthritic, osteopathic, ophthalmological and
 CC cyostatic activities. This polypeptide sequence is a human heart
 CC mitochondrial protein of the invention.

XX SQ Sequence 303 AA;

Query Match 65.5%; Score 19; DB 7; Length 303;
 Best Local Similarity 100.0%; Pred. No. 1.2e-10;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RAAFLKENTALTREVAEL 19

Db 264 RAAFLKENTALTREVAEL 282

RESULT 5

ADJ82509

ID ADJ82509 standard; protein; 303 AA.

XX AC ADJ82509;

XX DT 22-APR-2004 (first entry)

XX DE Human modifier of p21 (MP21) protein sequence SeqID75.

XX KW p21 pathway modulating agent; assay system; MP21; cyostatic;

XX KW MP21 protein activity modulator; cancer; genetically modified animal;

XX KW human.

XX OS Homo sapiens.

XX PN WO2004005486-A2.

XX PD 15-JAN-2004.

XX PF 09-JUL-2003; 2003WO-US021510.

XX PR 10-JUL-2002; 2002US-0394795P.

XX PR 07-AUG-2002; 2002US-0401739P.

XX PR 16-SEP-2002; 2002US-0411010P.

XX PR 30-DEC-2002; 2002US-0437158P.

XX (EXEL-) EXELIXIS INC.

XX PI Francis-Lang H, Friedman L, Kidd T, Roche S, Joo DM, Lickteig K;

XX PI Amundsen CD, Hai B, Zhang H, Adamkewicz JJ, Hammonds GR;

XX DR WPI; 2004-091358/09.

XX DR N-PSDB; ADI82448.

XX PT Identifying a candidate p21 pathway modulating agent, useful for treating
 PT a disease such as cancer, comprises contacting an assay system comprising
 PT a MP21 polypeptide or nucleic acid with a test agent.

XX PS Example 2; SEQ ID NO 75; 392pp; English.

XX CC This invention relates to a novel candidate p21 pathway modulating agent
 CC by contacting an assay system comprising an MP21 (modifier of p21)
 CC polypeptide or nucleic acid with a test agent, where in the absence of
 CC the test agent the system provides a reference activity and detecting a

CC test agent-biased activity of the assay system. The invention may be
 CC useful for the production of compounds with a cytostatic activity through
 CC modulation of MP21 protein activity. The MP21 polypeptide or nucleic acid
 CC can be used for identifying MP21 modulating agents useful as therapeutic
 CC targets for diagnosing cancer or treating disorders associated with
 CC defective or impaired p21 and/or MP21 function. MP21 modulating agents
 CC are useful in diagnosis, therapy, for example treating cancer, and
 CC pharmaceutical development. The genetically modified animals may be used
 CC for in vivo assays to test for activity of a candidate p21 modulating
 CC agent, or to further assess the role of MP21 in a p21 pathway process.
 CC The present sequence is that of a human MP21 protein which is an
 CC orthologue of a Drosophila p21 modifier and which was used in the
 CC exemplification of the invention.
 CC
 XX SQ Sequence 303 AA;
 Query Match 65.5%; Score 19; DB 8; Length 303;
 Best Local Similarity 100.0%; Pred. No. 1.2e-10;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 RAAFLEKENTALTREVAEL 19
 DB 264 RAAFLEKENTALTREVAEL 282
 RESULT 6
 ID ABM82367 standard; protein; 303 AA.
 XX
 AC ABM82367;
 XX
 DT 18-NOV-2004 (first entry)
 XX
 DE Tumour-associated antigenic target (TAT) polypeptide PRO83291, SEQ:6078.
 XX
 KW Tumour-associated antigenic target; TAT; human; overexpression; cancer;
 KW tumour; diagnosis; cell proliferative disorder; breast cancer;
 KW colorectal cancer; lung cancer; ovarian cancer; liver cancer;
 KW central nervous system cancer; bladder cancer; pancreatic cancer;
 KW cervical cancer; melanoma; leukaemia; hybridisation probe;
 KW chromosome identification; chromosome mapping; gene mapping;
 KW gene therapy; cytostatic.
 XX
 OS Homo sapiens.
 XX
 XX WO2004030615-A2.
 XX
 XX 15-APR-2004.
 XX
 XX 29-SEP-2003; 2003WO-US028547.
 XX
 XX 02-OCT-2002; 2002US-0414971P.
 XX
 XX (GETH) GENENTECH INC.
 XX
 XX Wu TD, Zhang Z, Zhou Y;
 XX
 XX WPI; 2004-347921/32.
 DR
 DR N-PSDB; ACN40954.
 XX
 XX New tumor-associated antigenic target polypeptides and nucleic acids,
 PT useful in preparing a medicament for treating or detecting a
 PT proliferative disorder, e.g. breast, lung, colorectal, ovarian or
 PT prostate cancer or tumor.
 XX
 PS Claim 12; SEQ ID NO 6078; 7273pp; English.
 XX
 XX The invention relates to human tumour-associated antigenic target (TAT)
 CC polypeptides, and their related nucleic acids. The TAT polypeptides are
 CC overexpressed in cancer tissues compared to normal tissues, and may thus
 CC serve as effective targets for the diagnosis and treatment of cancer in
 CC mammals. The invention also relates to nucleic acid and polypeptide
 CC sequences at least 80% identical to the TAT nucleic acids and

CC polypeptides; expression vectors and host cells comprising a TAT nucleic
 CC acid; an antibody specific for a TAT polypeptide; a peptide or organic
 CC molecule which binds to a TAT polypeptide; fusion proteins comprising a
 CC TAT polypeptide; and methods and compositions for the treatment or
 CC diagnosis of cancer in mammals. TAT polypeptides, nucleic acids,
 CC antibodies, antagonists, binding molecules and compositions are useful
 CC for diagnosing or treating a cell proliferative disorder associated with
 CC increased TAT expression, particularly cancers such as breast cancer,
 CC colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder
 CC cancer, pancreatic cancer, cervical cancer, cancers of the central
 CC nervous system, melanoma and leukaemia. TAT nucleic acids may further be
 CC used as hybridisation probes, in chromosome and gene mapping, in
 CC chromosome identification and in gene therapy. The present sequence
 CC represents a TAT polypeptide of the invention
 XX
 XX SQ Sequence 303 AA;
 Query Match 65.5%; Score 19; DB 8; Length 303;
 Best Local Similarity 100.0%; Pred. No. 1.2e-10;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 RAAFLEKENTALTREVAEL 19
 DB 264 RAAFLEKENTALTREVAEL 282
 RESULT 7
 ID ADI34277 standard; protein; 29 AA.
 XX
 AC ADI34277;
 XX
 DT 15-APR-2004 (first entry)
 XX
 DE Zipper peptide #2 for cross linking adenoviral ligands.
 XX
 KW adenovirus vector; Cytostatic; HER2/neu; CD40; tumor necrosis factor;
 KW TNF.
 XX
 OS Unidentified.
 XX
 XX WO2004009133-A1.
 XX
 XX 29-JAN-2004.
 XX
 XX 22-JUL-2003; 2003WO-US022852.
 XX
 XX 22-JUL-2002; 2002US-0397951P.
 XX
 XX (VECT-) VECTORLOGICS INC.
 XX
 XX Korokhov N, Mikheeva G;
 XX
 XX WPI; 2004-132871/13.
 XX
 XX Novel recombinant adenovirus having fiber protein modified by insertion
 PT of first zipper peptide that can crosslink to second zipper peptide-
 PT targeting ligand fusion protein, and binding between zipper peptides
 PT targets vector to cell.
 XX
 PS Claim 2; SEQ ID NO 2; 54pp; English.
 XX
 XX The present invention relates to a targeted recombinant adenovirus
 CC vector. The invention is useful for expressing a heterologous protein
 CC chosen from a tumor associated antigen, HER2/neu and carcinoembryonic
 CC antigen, in a target cell e.g., CD40 + cells, where the targeted
 CC useful in a method of gene transfer to CD40 + cells, where the targeted
 CC adenovirus vector mediates transfer of the gene encoding heterologous
 CC protein to the cell such as the dendritic cell. The vector is useful in
 CC gene therapy techniques for treatment of tumors. Multivalent interaction
 CC or trimeric CD40L with CD40 receptors causes CD40 ligation, which then
 CC results in enhanced survival of these cells and secretion of cytokines
 CC such as interleukin (IL)-1, IL-6, IL-8, IL-10, IL-12, tumor necrosis

CC factor (TNF)-4sagr;; macrophage inflammatory protein (MIP)-1a and enzymes
 CC such as matrix metalloproteinase. CD40-CD40L interaction also enhances
 CC monocyte tumoricidal activity. The present sequence represents a zipper
 CC peptide to be used for cross-linking targeting ligands with adenoviral
 CC virions.
 XX
 SQ Sequence 29 AA;

Query Match 48.3%; Score 14; DB 8; Length 29;
 Best Local Similarity 100.0%; Pred. NO. 1.3e-06;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 RAAFLEKENTALT 14
 DB 1 RAAFLEKENTALT 14
 |||||

RESULT 8
 ADI34279
 ID ADI34279 standard; protein; 43 AA.
 XX AC
 XX ADI34279;
 XX
 DT 15-APR-2004 (first entry)
 XX
 DE Zipper peptide #4 for cross linking adenoviral ligands.
 XX
 KW adenovirus vector; Cytostatic; HER2/neu; CD40; tumor necrosis factor;
 KW TNF.
 XX
 OS Unidentified.
 XX
 PN WO2004009133-A1.
 XX
 PD 29-JAN-2004.
 XX
 PF 22-JUL-2003; 2003WO-US022852.
 XX
 PR 22-JUL-2002; 2002US-037951P.
 XX
 PA (VECT-) VECTORLOGICS INC.
 XX
 PI Korokhov N, Mikheeva G;
 XX
 XX WPI; 2004-132871/13.

Novel recombinant adenovirus having fiber protein modified by insertion
 of first zipper peptide that can crosslink to second zipper peptide-
 targeting ligand fusion protein, and binding between zipper peptides
 targets vector to cell.

Claim 2; SEQ ID NO 4; 54pp; English.
 The present invention relates to a targeted recombinant adenovirus
 vector. The invention is useful for expressing a heterologous protein
 chosen from a tumor associated antigen, HER2/neu and carcinoembryonic
 antigen, in a target cell e.g., CD40 + cell such as dendritic cells, is
 useful in a method of gene transfer to CD40 + cells, where the targeted
 adenovirus vector mediates transfer of the gene encoding heterologous
 protein to the cell such as the dendritic cell. The vector is useful in
 gene therapy techniques for treatment of tumors. Multivalent interaction
 or trimeric CD40L with CD40 receptors causes CD40 ligation, which then
 results in enhanced survival of these cells and secretion of cytokines
 such as interleukin (IL)-1, IL-6, IL-8, IL-10, IL-12, tumor necrosis
 factor (TNF)-4sagr;; macrophage inflammatory protein (MIP)-1a and enzymes
 such as matrix metalloproteinase. CD40-CD40L interaction also enhances
 monocyte tumoricidal activity. The present sequence represents a zipper
 peptide to be used for cross-linking targeting ligands with adenoviral
 virions.

Sequence 43 AA;
 Query Match 41.4%; Score 12; DB 8; Length 43;

Best Local Similarity 100.0%; Pred. No. 0.00017;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 9 NTALRTEVAELE 20
 DB 12 NTALRTEVAELE 23
 |||||

RESULT 9
 ADE56596
 ID ADE56596 standard; protein; 325 AA.
 XX AC ADE56596;
 XX
 DT 29-JAN-2004 (first entry)
 XX
 DE Human Protein Q10586, SEQ ID NO 2450.
 XX
 KW Human; pain; neuronal tissue; gene therapy;
 KW spinal segmental nerve injury; chronic constriction injury; CCI;
 KW spared nerve injury; SNI; Chung.
 XX
 OS Homo sapiens.
 XX
 PN WO2003016475-A2.
 XX
 PD 27-FEB-2003.
 XX
 PF 14-AUG-2002; 2002WO-US025765.
 XX
 PR 14-AUG-2001; 2001US-0312147P.
 PR 01-NOV-2001; 2001US-0346382P.
 PR 26-NOV-2001; 2001US-0333347P.
 XX
 PA (GEO) GEN HOSPITAL CORP.
 PA (PABB) BAYER AG.
 XX
 PI Woolf C, D'urso D, Befort K, Costigan M;
 XX
 XX WPI; 2003-268312/26.
 DR GENEANK; Q10586.
 XX

New composition comprising two or more isolated polypeptides, useful for
 preparing a medicament for treating pain in an animal.
 Claim 1; Page; 1017pp; English.

The invention discloses a composition comprising two or more isolated rat
 or human polynucleotides or a polynucleotide which represents a fragment,
 derivative or allelic variation of the nucleic acid sequence. Also
 claimed are a vector comprising the novel polynucleotide, a host cell
 comprising the vector, a method for identifying a nucleotide sequence
 which is differentially regulated in an animal subjected to pain and a
 kit to perform the method, an array, a method for identifying an agent
 that increases or decreases the expression of the polynucleotide sequence
 that is differentially expressed in neuronal tissue of a first animal
 subjected to pain, a method for identifying a compound which regulates
 the expression of a polynucleotide sequence which is differentially
 expressed in an animal subjected to pain, a method for identifying a
 compound that regulates the activity of one or more of the
 polynucleotides, a method for producing a pharmaceutical composition, a
 method for identifying a compound or small molecule that regulates the
 activity in an animal of one or more of the polypeptides given in the
 specification, a method for identifying a compound useful in treating
 pain and a pharmaceutical composition comprising the one or more
 polypeptides or their antibodies. The polynucleotide or the compound that
 modulates its activity is useful for preparing a medicament for treating
 pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
 injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
 therapy). The sequence presented is a human protein (shown in Table 2 of
 the specification) which is differentially expressed during pain. Note:
 The sequence data for this patent did not form part of the printed
 specification, but was obtained in electronic form directly from WIPO at

```
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 325 AA;
    Query Match      31.0%; Score 9; DB 7; Length 325;
    Best Local Similarity 100.0%; Pred. No. 0.88;
    Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RAAFLEKEN 9
   |||||
Db 286 RAAFLEKEN 294

RESULT 10
ADD46912
ID ADD46912 standard; protein; 325 AA.
XX
AC ADD46912;
XX
XX 02-DEC-2004 (revised)
DT 29-JAN-2004 (first entry)
XX
DE Rat Protein AAA41083, SEQ ID NO 12598.
XX
XX Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
XX
OS Rattus norvegicus.
OS Unidentified.
XX
FN WO2003016475-A2.
XX
PD 27-FEB-2003.
XX
PF 14-AUG-2002; 2002WO-US025765.
XX
PR 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-0333347P.
XX
XX (GEHO ) GEN HOSPITAL CORP.
PA (FARB ) BAYER AG.
XX
XX Woolf C, D'urso D, Befort K, Costigan M;
XX
XX WPI; 2003-268312/26.
XX GENBANK; AAA41083.
XX
XX New composition comprising two or more isolated polypeptides, useful for
XX preparing a medicament for treating pain in an animal.
XX
XX Example 1; Page; 1017pp; English.
XX
XX The invention discloses a composition comprising two or more isolated rat
XX or human polynucleotides or a polynucleotide which represents a fragment,
XX derivative or allelic variation of the nucleic acid sequence. Also
XX claimed are a vector comprising the novel polynucleotide, a host cell
XX comprising the vector, a method for identifying a nucleotide sequence
XX which is differentially regulated in an animal subjected to pain and a
XX kit to perform the method, an array, a method for identifying an agent
XX that increases or decreases the expression of the polynucleotide sequence
XX that is differentially expressed in neuronal tissue of a first animal
XX subjected to pain, a method for identifying a compound which regulates
XX the expression of a polynucleotide sequence which is differentially
XX expressed in an animal subjected to pain, a method for identifying a
XX compound that regulates the activity of one or more of the
XX polynucleotides, a method for producing a pharmaceutical composition, a
XX method for identifying a compound or small molecule that regulates the
XX activity in an animal of one or more of the polypeptides given in the
XX pain and a pharmaceutical composition comprising the one or more
XX polypeptides or their antibodies. The polynucleotide or the compound that
XX modulates its activity is useful for preparing a medicament for treating
```

```
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a rat protein (described in Table 3
CC of the specification) which is differentially expressed during pain.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 325 AA;
    Query Match      31.0%; Score 9; DB 7; Length 325;
    Best Local Similarity 100.0%; Pred. No. 0.88;
    Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RAAFLEKEN 9
   |||||
Db 286 RAAFLEKEN 294

RESULT 11
ADE56594
ID ADE56594 standard; protein; 325 AA.
XX
AC ADE56594;
XX
XX 29-JAN-2004 (first entry)
DT
XX
DE Rat Protein P16443, SEQ ID NO 2448.
XX
XX Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
XX
OS Rattus norvegicus.
XX
XX WO2003016475-A2.
XX
XX 27-FEB-2003.
XX
XX 14-AUG-2002; 2002WO-US025765.
XX
XX 14-AUG-2001; 2001US-0312147P.
XX 01-NOV-2001; 2001US-0346382P.
XX 26-NOV-2001; 2001US-0333347P.
XX
XX (GEHO ) GEN HOSPITAL CORP.
PA (FARB ) BAYER AG.
XX
XX Woolf C, D'urso D, Befort K, Costigan M;
XX
XX WPI; 2003-268312/26.
XX GENBANK; P16443.
XX
XX New composition comprising two or more isolated polypeptides, useful for
XX preparing a medicament for treating pain in an animal.
XX
XX Claim 1; Page; 1017pp; English.
XX
XX The invention discloses a composition comprising two or more isolated rat
XX or human polynucleotides or a polynucleotide which represents a fragment,
XX derivative or allelic variation of the nucleic acid sequence. Also
XX claimed are a vector comprising the novel polynucleotide, a host cell
XX comprising the vector, a method for identifying a nucleotide sequence
XX which is differentially regulated in an animal subjected to pain and a
XX kit to perform the method, an array, a method for identifying an agent
XX that increases or decreases the expression of the polynucleotide sequence
XX that is differentially expressed in neuronal tissue of a first animal
XX subjected to pain, a method for identifying a compound which regulates
XX the expression of a polynucleotide sequence which is differentially
XX expressed in an animal subjected to pain, a method for identifying a
XX compound that regulates the activity of one or more of the
XX polynucleotides, a method for producing a pharmaceutical composition, a
XX method for identifying a compound or small molecule that regulates the
XX activity in an animal of one or more of the polypeptides given in the
XX pain and a pharmaceutical composition comprising the one or more
XX polypeptides or their antibodies. The polynucleotide or the compound that
XX modulates its activity is useful for preparing a medicament for treating
```

CC specification, a method for identifying a compound useful in treating
 CC pain and a pharmaceutical composition comprising the one or more
 CC polypeptides or their antibodies. The polynucleotide or the compound that
 CC modulates its activity is useful for preparing a medicament for treating
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
 CC therapy). The sequence presented is a rat protein (shown in Table 2 of
 CC the specification) which is differentially expressed during pain. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic form directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 325 AA;
 Query Match 31.0%; Score 9; DB 7; Length 325;
 Best Local Similarity 100.0%; Pred. No. 0.88;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RAAFLKEN 9
 DB 286 RAAFLKEN 294
 ADD46916
 ID ADD46916 standard; protein; 325 AA.
 AC ADD46916;

RESULT 12
 ADD46916
 ID ADD46916 standard; protein; 325 AA.
 XX
 AC ADD46916;
 DT 02-DEC-2004 (revised)
 DT 29-JAN-2004 (first entry)

DE Rat Protein AAA41083, SEQ ID NO 12602.
 KW Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
 KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.

OS Rattus norvegicus.
 OS Unidentified.
 XX WO2003016475-A2.
 XX 27-FEB-2003.

XX 14-AUG-2002; 2002WO-US025765.
 XX 14-AUG-2001; 2001US-0312147P.
 PR 01-NOV-2001; 2001US-0346382P.
 PR 26-NOV-2001; 2001US-0333347P.

XX (GEO) GEN HOSPITAL CORP.
 PA (FARB) BAYER AG.

XX Woolf C, D'urso D, Befort K, Costigan M;
 XX WPI; 2003-268312/26.
 DR GENBANK; AAA41083.

XX New composition comprising two or more isolated polypeptides, useful for
 PT preparing a medicament for treating pain in an animal.

XX Example 1; Page; 1017pp; English.

XX The invention discloses a composition comprising two or more isolated rat
 CC or human polynucleotides or a polynucleotide which represents a fragment,
 CC derivative or allelic variation of the nucleic acid sequence. Also
 CC claimed are a vector comprising the novel polynucleotide, a host cell
 CC comprising the vector, a method for identifying a nucleotide sequence
 CC which is differentially regulated in an animal subjected to pain and a
 CC kit to perform the method, an array, a method for identifying an agent
 CC that increases or decreases the expression of the polynucleotide sequence
 CC that is differentially expressed in neuronal tissue of a first animal
 CC subjected to pain, a method for identifying a compound which regulates

CC the expression of a polynucleotide sequence which is differentially
 CC expressed in an animal subjected to pain, a method for identifying a
 CC compound that regulates the activity of one or more of the
 CC polynucleotides, a method for producing a pharmaceutical composition, a
 CC method for identifying a compound or small molecule that regulates the
 CC activity in an animal of one or more of the polypeptides given in the
 CC specification, a method for identifying a compound useful in treating
 CC pain and a pharmaceutical composition comprising the one or more
 CC polypeptides or their antibodies. The polynucleotide or the compound that
 CC modulates its activity is useful for preparing a medicament for treating
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
 CC therapy). The sequence presented is a rat protein (described in Table 3
 CC of the specification) which is differentially expressed during pain.
 CC Note: the sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic form directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 325 AA;
 Query Match 31.0%; Score 9; DB 7; Length 325;
 Best Local Similarity 100.0%; Pred. No. 0.88;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RAAFLKEN 9
 DB 286 RAAFLKEN 294
 ADD46914
 ID ADD46914 standard; protein; 325 AA.

XX
 AC ADD46914;
 DT 02-DEC-2004 (revised)
 DT 29-JAN-2004 (first entry)

XX Human Protein NP_001343, SEQ ID NO 12600.
 XX Human; pain; neuronal tissue; gene therapy;
 KW spinal segmental nerve injury; chronic constriction injury; CCI;
 KW spared nerve injury; SNI; Chung.

XX Homo sapiens.
 OS Unidentified.
 XX WO2003016475-A2.
 XX 27-FEB-2003.

XX 14-AUG-2002; 2002WO-US025765.
 XX 14-AUG-2001; 2001US-0312147P.
 PR 01-NOV-2001; 2001US-0346382P.
 PR 26-NOV-2001; 2001US-0333347P.

XX (GEO) GEN HOSPITAL CORP.
 PA (FARB) BAYER AG.

XX Woolf C, D'urso D, Befort K, Costigan M;
 XX WPI; 2003-268312/26.
 DR GENBANK; NP_001343.

XX New composition comprising two or more isolated polypeptides, useful for
 PT preparing a medicament for treating pain in an animal.

XX Example 1; Page; 1017pp; English.

XX The invention discloses a composition comprising two or more isolated rat
 CC or human polynucleotides or a polynucleotide which represents a fragment,
 CC derivative or allelic variation of the nucleic acid sequence. Also
 CC claimed are a vector comprising the novel polynucleotide, a host cell
 CC comprising the vector, a method for identifying a nucleotide sequence
 CC which is differentially regulated in an animal subjected to pain and a
 CC kit to perform the method, an array, a method for identifying an agent
 CC that increases or decreases the expression of the polynucleotide sequence
 CC that is differentially expressed in neuronal tissue of a first animal
 CC subjected to pain, a method for identifying a compound which regulates

CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a human protein (described in Table 3
CC of the specification) which is differentially expressed during pain.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 325 AA;

Query Match 31.0%; Score 9; DB 7; Length 325;

Best Local Similarity 100.0%; Pred. No. 0.88;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RAAFLXEN 9

|||||||

286 RAAFLXEN 294

RESULT 14

ID ADD46918 standard; protein; 325 AA.

XX AC ADD46918;

XX DT 02-DEC-2004 (revised)

XX DT 29-JAN-2004 (first entry)

XX DE Human Protein NP_001343, SEQ ID NO 12604.

XX KW Human; pain; neuronal tissue; gene therapy;
XX KW spinal segmental nerve injury; chronic constriction injury; CCI;
XX KW spared nerve injury; SNI; Chung.

XX OS Homo sapiens.

XX OS Unidentified.

XX FN WO2003016475-A2.

XX PD 27-FEB-2003.

XX PF 14-AUG-2002; 2002WO-US025765.

XX PR 14-AUG-2001; 2001US-0312147P.

XX PR 01-NOV-2001; 2001US-0346382P.

XX PR 26-NOV-2001; 2001US-0333347P.

XX XX (GEO) GEN HOSPITAL CORP.

XX PA (FARB) BAYER AG.

XX XX Woolf C, D'urso D, Befort K, Costigan M;

XX XX WPI; 2003-268312/26.

XX DR GENBANK; NP_001343.

XX XX New composition comprising two or more isolated polypeptides, useful for

PT preparing a medicament for treating pain in an animal.

XX Example 1; Page; 1017pp; English.

XX The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a human protein (described in Table 3
CC of the specification) which is differentially expressed during pain.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 325 AA;

Query Match 31.0%; Score 9; DB 7; Length 325;

Best Local Similarity 100.0%; Pred. No. 0.88;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RAAFLXEN 9

|||||||

286 RAAFLXEN 294

RESULT 15

ID ADE56590

XX AC ADE56590 standard; protein; 325 AA.

XX AC ADE56590;

XX DT 29-JAN-2004 (first entry)

XX DE Rat Protein P16443, SEQ ID NO 2444.

XX KW Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
XX KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.

XX OS Rattus norvegicus.

XX FN WO2003016475-A2.

XX PD 27-FEB-2003.

XX PF 14-AUG-2002; 2002WO-US025765.

XX PR 14-AUG-2001; 2001US-0312147P.

XX PR 01-NOV-2001; 2001US-0346382P.

XX PR 26-NOV-2001; 2001US-0333347P.

XX XX (GEO) GEN HOSPITAL CORP.

XX PA (FARB) BAYER AG.

XX PI Woolf C, D'urso D, Befort K, Costigan M;

XX XX WPI; 2003-268312/26.

XX DR GENBANK; NP_001343.

XX XX New composition comprising two or more isolated polypeptides, useful for

DR WPI; 2003-269312/26.
XX GENBANK; P16443.

PT New composition comprising two or more isolated polypeptides, useful for
XX preparing a medicament for treating pain in an animal.

PS Claim 1; Page; 1017pp; English.

XX The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (SN1), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a rat protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 325 AA;

Query Match 31.0%; Score 9; DB 7; Length 325;
Best Local Similarity 100.0%; Pred. No. 0.88;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RAAFLKEN 9
Db 286 RAAFLKEN 294

Search completed: July 26, 2005, 12:14:04
Job time : 76.3055 secs

Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 26, 2005, 12:02:34 ; Search time 19.5347 Seconds
(without alignments)
110.819 Million cell updates/sec

Title: us-10-624-218-1

Perfect score: 29

Sequence: 1 RAAFLKENTALTREVALEKEVGRCNI 29

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 513545 seqs, 74649064 residues

Word size : 0

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents_AA.*

- 1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	19	65.5	345	4	US-09-949-016-10895
2	9	31.0	335	4	Sequence 10895, A
3	8	27.6	334	6	Sequence 11210, A
4	8	27.6	334	6	Patent No. 5290690
5	8	27.6	534	4	Patent No. 5290690-11
6	7	24.1	84	4	Sequence 2, Appli
7	7	24.1	431	4	Sequence 16607, A
8	6	20.7	95	4	Sequence 5352, Ap
9	6	20.7	105	4	Sequence 4267, Ap
10	6	20.7	112	3	Sequence 5144, Ap
11	6	20.7	121	4	Sequence 75, Appl
12	6	20.7	124	2	Sequence 5995, Ap
13	6	20.7	132	4	Sequence 12, Appl
14	6	20.7	135	4	Sequence 10568, A
15	6	20.7	141	4	Sequence 5156, Ap
16	6	20.7	142	4	Sequence 3869, Ap
17	6	20.7	148	4	Sequence 6345, Ap
18	6	20.7	149	4	Sequence 13147, A
19	6	20.7	149	4	Sequence 842, App
20	6	20.7	164	4	Sequence 855, App
21	6	20.7	176	4	Sequence 94, Appl
22	6	20.7	178	4	Sequence 16790, A
23	6	20.7	181	4	Sequence 4001, Ap
24	6	20.7	185	2	Sequence 16651, A
25	6	20.7	185	3	Sequence 2, Appli
26	6	20.7	185	3	Sequence 1, Appli
27	6	20.7	185	3	Sequence 1, Appli

Sequence 6, Appli
Sequence 12, Appl
Sequence 36, Appl
Sequence 4068, Ap
Sequence 5683, Ap
Sequence 18199, A
Sequence 16060, A
Sequence 15, Appl
Sequence 16, Appl
Sequence 9954, Ap
Sequence 2, Appli
Sequence 49, Appl
Sequence 6138, Ap
Sequence 10615, A
Sequence 16801, A
Sequence 16151, A
Sequence 3064, Ap

28 6 20.7 185 4 US-09-216-430C-6
29 6 20.7 197 3 US-09-370-473-12
30 6 20.7 198 4 US-08-529-055-36
31 6 20.7 198 4 US-09-107-433-4068
32 6 20.7 267 4 US-09-107-532A-5683
33 6 20.7 301 4 US-09-248-796A-18199
34 6 20.7 304 4 US-09-248-796A-16060
35 6 20.7 315 4 US-09-674-529B-15
36 6 20.7 316 4 US-09-674-529B-16
37 6 20.7 316 4 US-09-949-016-9954
38 6 20.7 319 3 US-09-347-803-2
39 6 20.7 325 3 US-09-108-020-49
40 6 20.7 325 4 US-09-685-296-49
41 6 20.7 333 4 US-09-949-016-6128
42 6 20.7 340 4 US-09-949-016-10615
43 6 20.7 344 4 US-09-902-540-16801
44 6 20.7 361 4 US-09-902-540-16151
45 6 20.7 362 4 US-09-583-110-3064

ALIGNMENTS

RESULT 1
US-09-949-016-10895
; Sequence 10895, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10895
; LENGTH: 345
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-10895

Query Match 65.5%; Score 19; DB 4; Length 345;
Best Local Similarity 100.0%; Pred. No. 1.8e-12;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RAAFLKENTALTREVAEL 19
DB 306 RAAFLKENTALTREVAEL 324

RESULT 2
US-09-949-016-11210
; Sequence 11210, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498

us-10-624-218-1.olg.ra1

Thu Jul 28 11:34:39 2005

```
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 11210
; LENGTH: 335
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-11210

Query Match      31.0%; Score 9; DB 4; Length 335;
Best Local Similarity 100.0%; Pred. No. 0.077;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 RAAFLEREN 9
Db      296 RAAFLEREN 304
|||||

RESULT 3
5290690-11
; Patent No. 5290690
; APPLICANT: MRABET, NADIR; LASTERS, IGNACE; STANSSENS, PATRICK
; MATTHYSSENS, GASTON; WODAK, SHOSHANA; QUAX, WILHELMUS J.
; TITLE OF INVENTION: METHODS AND MEANS FOR CONTROLLING THE
; STABILITY OF PROTEINS
; NUMBER OF SEQUENCES: 22
; CURRENT APPLICATION NUMBER: US/07/398,706
; FILING DATE: 25-AUG-1989
; SEQ ID NO: 11; Length: 334
; LENGTH: 334
5290690-11

Query Match      27.6%; Score 8; DB 6; Length 334;
Best Local Similarity 100.0%; Pred. No. 0.89;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      16 VAELEKEV 23
Db      244 VAELEKEV 251
|||||

RESULT 4
5290690-11
; Patent No. 5290690
; APPLICANT: MRABET, NADIR; LASTERS, IGNACE; STANSSENS, PATRICK
; MATTHYSSENS, GASTON; WODAK, SHOSHANA; QUAX, WILHELMUS J.
; TITLE OF INVENTION: METHODS AND MEANS FOR CONTROLLING THE
; STABILITY OF PROTEINS
; NUMBER OF SEQUENCES: 22
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/398,706
; FILING DATE: 25-AUG-1989
; SEQ ID NO: 11; Length: 334
; LENGTH: 334
5290690-11

Query Match      27.6%; Score 8; DB 6; Length 334;
Best Local Similarity 100.0%; Pred. No. 0.89;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      16 VAELEKEV 23
Db      244 VAELEKEV 251
|||||

Query Match      27.6%; Score 9; DB 4; Length 335;
Best Local Similarity 100.0%; Pred. No. 0.077;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 RAAFLEREN 9
Db      296 RAAFLEREN 304
|||||

; APPLICANT: Artiushin, Sergey
; TITLE OF INVENTION: Compounds Encoding the Protective M-Like Protein of Streptococcus
; TITLE OF INVENTION: and Assays Therefor
; FILE REFERENCE: 50229-212
; CURRENT APPLICATION NUMBER: US/09/103,664A
; CURRENT FILING DATE: 1998-06-23
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 534
; TYPE: PRT
; ORGANISM: Streptococcus equi
US-09-103-664A-2

Query Match      27.6%; Score 8; DB 4; Length 534;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      16 VAELEKEV 23
Db      391 VAELEKEV 398
|||||

RESULT 6
US-09-902-540-16607
; Sequence 16607, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 16607
; LENGTH: 84
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-16607

Query Match      24.1%; Score 7; DB 4; Length 84;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 RAAFLER 7
Db      70 RAAFLER 76
|||||

RESULT 7
US-09-134-000C-5352
; Sequence 5352, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5352
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
```


US-09-134-000C-5352

Query Match 24.1%; Score 7; DB 4; Length 431;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 VAELEKE 22
Db 151 VAELEKE 157

RESULT 8

US-09-621-976-4267
; Sequence 4267, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 4267
; LENGTH: 95
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-621-976-4267

Query Match 20.7%; Score 6; DB 4; Length 95;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 ELEKEV 23
Db 84 ELEKEV 89

RESULT 9

US-09-583-110-5144
; Sequence 5144, Application US/09583110
; Patent No. 6699703
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; FILE REFERENCE: PATH00-07A
; CURRENT APPLICATION NUMBER: US/09/583,110
; CURRENT FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 5144
; LENGTH: 105
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-583-110-5144

Query Match 20.7%; Score 6; DB 4; Length 105;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 ENTALR 13
Db 37 ENTALR 42

RESULT 10

US-08-857-076-75
; Sequence 75, Application US/08857076C
; Patent No. 6225120
; GENERAL INFORMATION:
; APPLICANT: Ruvkun, Gary
; APPLICANT: Kimura, Koutarou
; APPLICANT: Patterson, Garth
; APPLICANT: Ogg, Scott
; APPLICANT: Paradis, Suzanne
; APPLICANT: Tissenbaum, Heidi
; APPLICANT: Morris, Jason
; APPLICANT: Kowsek, Allison
; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC TOOLS FOR
; FILE REFERENCE: 00786/351001
; CURRENT APPLICATION NUMBER: US/08/857,076C
; CURRENT FILING DATE: 1997-05-15
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 75
; LENGTH: 112
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-08-857-076-75

Query Match 20.7%; Score 6; DB 3; Length 112;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 AELEKE 22
Db 28 AELEKE 33

RESULT 11

US-09-134-000C-5995
; Sequence 5995, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 5995
; LENGTH: 121
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (35)..(35)
; OTHER INFORMATION: Amino acid 35 is Xaa wherein Xaa = any amino acid.
US-09-134-000C-5995

Query Match 20.7%; Score 6; DB 4; Length 121;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 LEKEVG 24
Db 26 LEKEVG 31

RESULT 12

US-08-743-200-12
; Sequence 12, Application US/08743200
; Patent No. 5861260

Thu Jul 28 11:34:39 2005

GENERAL INFORMATION:
 APPLICANT: Doxsey, Stephen J.
 TITLE OF INVENTION: DIAGNOSTIC METHODS FOR SCREENING
 TITLE OF INVENTION: PATIENTS FOR SCLERODERMA
 NUMBER OF SEQUENCES: 36
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Fish & Richardson P.C.
 STREET: 225 Franklin Street
 CITY: Boston
 STATE: MA
 COUNTRY: US
 ZIP: 02110-2804
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSeq Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/743,200
 FILING DATE: 05-NOV-1996
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER:
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Fasse, J. Peter
 REGISTRATION NUMBER: 32,983
 REFERENCE/DOCKET NUMBER: 07917/025001
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617-542-5070
 TELEFAX: 617-542-8906
 INFORMATION FOR SEQ ID NO: 12:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 124 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 FRAGMENT TYPE: internal
 US-08-743-200-12

Query Match 20.7%; Score 6; DB 2; Length 124;
 Best Local Similarity 100.0%; Pred. No. 44;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 AELEKE 22
 Db 61 AELEKE 66

RESULT 13
 US-09-489-039A-10568
 Sequence 10568, Application US/09489039A
 Patent No. 6610836
 GENERAL INFORMATION:
 APPLICANT: Gary Breton et. al
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
 FILE REFERENCE: 2709.2004001
 CURRENT APPLICATION NUMBER: US/09/489,039A
 CURRENT FILING DATE: 2000-01-27
 PRIOR APPLICATION NUMBER: US 60/117,747
 PRIOR FILING DATE: 1999-01-29
 NUMBER OF SEQ ID NOS: 14342
 SEQ ID NO 10568
 LENGTH: 132
 TYPE: PRT
 ORGANISM: Klebsiella pneumoniae
 US-09-489-039A-10568

Query Match 20.7%; Score 6; DB 4; Length 132;
 Best Local Similarity 100.0%; Pred. No. 47;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 AELEKE 22
 Db 48 AELEKE 53
 RESULT 14
 US-09-107-433-5156
 Sequence 5156, Application US/09107433
 Patent No. 6800744
 GENERAL INFORMATION:
 APPLICANT: Lynn A Doucette-Stamm and David Bush
 TITLE OF INVENTION: SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE
 THERAPEUTICS
 NUMBER OF SEQUENCES: 5206
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: GENOME THERAPEUTICS CORPORATION
 STREET: 100 Beaver Street
 CITY: Waltham
 STATE: Massachusetts
 COUNTRY: USA
 ZIP: 02354
 COMPUTER READABLE FORM:
 MEDIUM TYPE: CD-ROM ISO9660
 COMPUTER: <Unknown>
 OPERATING SYSTEM: <Unknown>
 SOFTWARE: <Unknown>
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/107,433
 FILING DATE: 30-Jun-1998
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 60/ 085131
 FILING DATE: May 12, 1998
 APPLICATION NUMBER: 60/051553
 FILING DATE: July 2, 1997
 ATTORNEY/AGENT INFORMATION:
 NAME: Ariniello, Pamela Deneke
 REGISTRATION NUMBER: 40,489
 REFERENCE/DOCKET NUMBER: GTC-011
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (781)893-5007
 TELEFAX: (781)893-8277
 INFORMATION FOR SEQ ID NO: 5156:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 135 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 HYPOTHETICAL: YES
 ORIGINAL SOURCE:
 ORGANISM: Streptococcus pneumoniae
 FEATURE:
 NAME/KEY: misc feature
 LOCATION: (B) LOCATION 1...135
 SEQUENCE DESCRIPTION: SEQ ID NO: 5156:
 US-09-107-433-5156

Query Match 20.7%; Score 6; DB 4; Length 135;
 Best Local Similarity 100.0%; Pred. No. 48;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 ENTALR 13
 Db 67 ENTALR 72

RESULT 15
 US-09-583-110-3869
 Sequence 3869, Application US/09583110
 Patent No. 6699703
 GENERAL INFORMATION:
 APPLICANT: Lynn Doucette-Stamm et al.
 TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus

; TITLE OF INVENTION: Pneumoniae for Diagnostics and Therapeutics
; FILE REFERENCE: PATH00-07A
; CURRENT APPLICATION NUMBER: US/09/583,110
; CURRENT FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 3869
; LENGTH: 141
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-583-110-3869

Query Match 20.7%; Score 6; DB 4; Length 141;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 FLEKEN 9
Db 21 FLEKEN 26

Search completed: July 26, 2005, 12:22:30
Job time : 21.5347 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 26, 2005, 12:08:05 ; Search time 68.2708 Seconds
(without alignments)
165.236 Million cell updates/sec

Title: US-10-624-218-1

Perfect score: 29

Sequence: 1 RAAFLKENTALTREVAELEKEVGRNCNI 29

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 1741741 seqs, 388992284 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1741741

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published Applications AA:
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
19: /cgn2_6/ptodata/2/pubpaa/US11A_PUBCOMB.pep.*
20: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
21: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
22: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	29	100.0	29	16 US-10-624-218-1	Sequence 1, Appli
2	19	65.5	303	16 US-10-408-765A-507	Sequence 507, App
3	14	48.3	29	16 US-10-624-218-2	Sequence 2, Appli
4	12	41.4	43	16 US-10-624-218-4	Sequence 4, Appli
5	8	27.6	166	15 US-10-424-599-173831	Sequence 173831,
6	8	27.6	567	15 US-10-282-122A-71601	Sequence 71601, A
7	7	24.1	121	17 US-10-489-123-5	Sequence 5, Appli
8	7	24.1	146	16 US-10-767-701-43502	Sequence 43502, A
9	7	24.1	163	15 US-10-320-797-3020	Sequence 3020, Ap
10	7	24.1	215	16 US-10-739-930-10206	Sequence 10206, A
11	7	24.1	245	10 US-09-907-907A-45	Sequence 45, Appli

12	7	24.1	284	15	US-10-369-493-3657	Sequence 3657, Ap
13	7	24.1	286	15	US-10-424-599-250565	Sequence 250565,
14	7	24.1	290	15	US-10-424-599-229602	Sequence 229602,
15	7	24.1	350	16	US-10-437-963-108061	Sequence 108061,
16	7	24.1	357	15	US-10-282-122A-54529	Sequence 54529, A
17	7	24.1	484	16	US-10-437-963-163307	Sequence 163307,
18	7	24.1	515	16	US-10-437-963-141509	Sequence 141509,
19	7	24.1	581	16	US-10-437-963-125056	Sequence 125056,
20	7	24.1	593	16	US-10-437-963-137447	Sequence 137447,
21	7	24.1	656	16	US-10-437-963-159911	Sequence 159911,
22	7	24.1	710	16	US-10-437-963-123987	Sequence 123987,
23	7	24.1	725	16	US-10-437-963-114661	Sequence 114661,
24	7	24.1	747	16	US-10-437-963-186162	Sequence 186162,
25	7	24.1	799	16	US-10-437-963-182178	Sequence 182178,
26	7	24.1	1083	15	US-10-394-575-93	Sequence 93, Appli
27	7	24.1	1441	17	US-10-732-923-3352	Sequence 3352, Ap
28	6	20.7	21	17	US-10-489-123-1	Sequence 1, Appli
29	6	20.7	21	17	US-10-489-123-15	Sequence 15, Appli
30	6	20.7	40	15	US-10-424-599-255314	Sequence 255314,
31	6	20.7	40	17	US-10-489-123-14	Sequence 14, Appli
32	6	20.7	50	16	US-10-437-963-125332	Sequence 125332,
33	6	20.7	57	17	US-10-489-123-13	Sequence 13, Appli
34	6	20.7	62	16	US-10-767-701-48001	Sequence 48001, A
35	6	20.7	64	16	US-10-767-701-51798	Sequence 51798, A
36	6	20.7	71	16	US-10-425-115-230497	Sequence 230497,
37	6	20.7	71	16	US-10-425-115-232160	Sequence 232160,
38	6	20.7	72	17	US-10-489-123-12	Sequence 12, Appli
39	6	20.7	76	16	US-10-767-701-40996	Sequence 40996, A
40	6	20.7	76	16	US-10-425-115-367583	Sequence 367583,
41	6	20.7	78	14	US-10-091-572-230	Sequence 230, App
42	6	20.7	79	15	US-10-424-599-255597	Sequence 255597,
43	6	20.7	82	16	US-10-425-115-228483	Sequence 228483,
44	6	20.7	85	16	US-10-425-115-367528	Sequence 367528,
45	6	20.7	85	17	US-10-489-123-11	Sequence 11, Appli

ALIGNMENTS

RESULT 1
US-10-624-218-1
; Sequence 1, Application US/10624218
; Publication No. US20040171158A1
; GENERAL INFORMATION:
; APPLICANT: Korokhov, Nikolay
; APPLICANT: Mikheeva, Galina
; TITLE OF INVENTION: Adenoviral Vector Incorporating Zipper
; TITLE OF INVENTION: Peptide-Modified Fiber Protein and Uses Thereof
; FILE REFERENCE: D6463
; CURRENT APPLICATION NUMBER: US/10/624,218
; CURRENT FILING DATE: 2003-07-22
; PRIOR APPLICATION NUMBER: US 60/397,951
; PRIOR FILING DATE: 2002-07-22
; NUMBER OF SEQ ID NOS: 4
; SEQ ID NO 1
; LENGTH: 29
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: PEPTIDE
; OTHER INFORMATION: zipper peptide E E34
US-10-624-218-1

Query Match	100.0%	Score 29;	DB 16;	Length 29;
Best Local Similarity	100.0%	Pred. No. 1.2e-21;		
Matches	29;	Conservative	0;	Mismatches
			0;	Indels
			0;	Gaps
				0;
Qy	1	RAAFLEKENTALTREVAELEKEVGRNCNI	29	
Db	1	RAAFLEKENTALTREVAELEKEVGRNCNI	29	

RESULT 2

```

US-10-408-765A-507
; TITLE OF INVENTION: Adenoviral Vector Incorporating Zipper
; TITLE OF INVENTION: Peptide-Modified Fiber Protein and Uses Thereof
; FILE REFERENCE: D6463
; CURRENT APPLICATION NUMBER: US/10/624,218
; CURRENT FILING DATE: 2003-07-22
; PRIOR APPLICATION NUMBER: US 60/397,951
; PRIOR FILING DATE: 2002-07-22
; NUMBER OF SEQ ID NOS: 4
; SEQ ID NO 4
; LENGTH: 43
; TYPE: PRT
; ORGANISM: Artificial Sequence
; NAME/KEY: PEPTIDE
; OTHER INFORMATION: zipper peptide RR12EE345L
US-10-624-218-4
Query Match 41.4%; Score 12; DB 16; Length 43;
Best Local Similarity 100.0%; Pred. No. 0.00013;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 NTALRTEVAELE 20
DB 12 NTALRTEVAELE 23

RESULT 5
US-10-424-599-173831
; Sequence 173831, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 173831
; LENGTH: 166
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_127989C.1.pap
US-10-424-599-173831
Query Match 27.6%; Score 8; DB 15; Length 166;
Best Local Similarity 100.0%; Pred. No. 4.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 LRTEVAEL 19
DB 112 LRTEVAEL 119

RESULT 6
US-10-282-122A-71601
; Sequence 71601, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert

```

```

; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 71601
; LENGTH: 567
; TYPE: PRT
; ORGANISM: Staphylococcus haemolyticus
US-10-282-122A-71601

Query Match      27.6%; Score 8; DB 15; Length 567;
Best Local Similarity 100.0%; Pred. No. 13;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      14 TEVAELEK 21
      |||||
DB      248 TEVAELEK 255

RESULT 7
US-10-489-123-5
; Sequence 5, Application US/10489123
; Publication No. US20050074463A1
; GENERAL INFORMATION:
; APPLICANT: AUTRAN, BRIGITTE
; APPLICANT: SAMRI, ASSIA
; APPLICANT: DEBRE, PATRICE
; APPLICANT: CALVEZ, VINCENT
; APPLICANT: KATLAMA, CHRISTINE
; APPLICANT: HAAS, GABY
; TITLE OF INVENTION: THERAPEUTIC VACCINATION METHOD, MUTATED PEPTIDES OF HIV
; TITLE OF INVENTION: REVERSE TRANSCRIPTASE AND THEIR USE FOR VACCINATION AND
; TITLE OF INVENTION: DIAGNOSTIC PURPOSES
; FILE REFERENCE: BDAC:007US
; CURRENT APPLICATION NUMBER: US/10/489,123
; CURRENT FILING DATE: 2004-03-12
; PRIOR APPLICATION NUMBER: PCT/FR01/02872
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 121
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-10-489-123-5

```

```

Query Match      24.1%; Score 7; DB 17; Length 121;
Best Local Similarity 100.0%; Pred. No. 30;
Matches      7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      17 AELEKEV 23
      |||||
DB      50 AELEKEV 56

RESULT 8
US-10-767-701-43502
; Sequence 43502, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 43502
; LENGTH: 146
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C45164_1.pep
US-10-767-701-43502

```

```

Query Match      24.1%; Score 7; DB 16; Length 146;
Best Local Similarity 100.0%; Pred. No. 36;
Matches      7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      16 VAELEKE 22
      |||||
DB      54 VAELEKE 60

```

```

RESULT 9
US-10-320-797-3020
; Sequence 3020, Application US/10320797
; Publication No. US20040014955A1
; GENERAL INFORMATION:
; APPLICANT: Eroshkin, Alexey M.
; APPLICANT: Zamudio, Carlos
; TITLE OF INVENTION: IDENTIFICATION OF ESSENTIAL GENES OF CRYPTOCOCCUS NEOFORMANS AND
; TITLE OF INVENTION: METHODS OF USE
; FILE REFERENCE: 10182-021-999
; CURRENT APPLICATION NUMBER: US/10/320,797
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: 60/341,261
; PRIOR FILING DATE: 2001-12-17
; NUMBER OF SEQ ID NOS: 3361
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3020
; LENGTH: 163
; TYPE: PRT
; ORGANISM: Cryptococcus neoformans
US-10-320-797-3020

```

```

Query Match      24.1%; Score 7; DB 15; Length 163;
Best Local Similarity 100.0%; Pred. No. 40;
Matches      7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      15 EVAELEK 21
      |||||
DB      44 EVAELEK 50

```

RESULT 10

```
US-10-739-930-10206
; Sequence 10206, Application US/10739930
; Publication No. US20040216190A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
; TITLE OF INVENTION: PLANTS AND USES THEREOF FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53377)B
; CURRENT APPLICATION NUMBER: US/10739,930
; CURRENT FILING DATE: 2003-12-18
; NUMBER OF SEQ ID NOS: 11088
; SEQ ID NO 10206
; LENGTH: 215
; TYPE: PRT
; ORGANISM: Triticum aestivum
; FEATURE:
; LOCATION: (1)..(284)
; OTHER INFORMATION: unsure at all xaa locations
US-10-739-930-10206
Query Match 24.1%; Score 7; DB 16; Length 215;
Best Local Similarity 100.0%; Pred. No. 51; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 AELEKEV 23
DB 58 AELEKEV 64

RESULT 11
US-09-907-907A-45
; Sequence 45, Application US/09907907A
; Publication No. US20030099660A1
; GENERAL INFORMATION:
; APPLICANT: Fisher, Paul B.
; APPLICANT: Leszczyniecka, Magdalena
; TITLE OF INVENTION: GENES DISPLAYING ENHANCED EXPRESSION DURING CELLULAR SENESECE A
; TITLE OF INVENTION: TERMINAL CELL DIFFERENTIATION AND USES THEREOF
; FILE REFERENCE: A34584-A-PCT-USA (070050.1664)
; CURRENT APPLICATION NUMBER: US/09/907,907A
; CURRENT FILING DATE: 2001-07-16
; PRIOR APPLICATION NUMBER: US 09/243,277
; PRIOR FILING DATE: 1999-02-02
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 45
; LENGTH: 245
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Consensus sequence between Homo sapiens OLD-35 and
; OTHER INFORMATION: Bacillus subtilis PNPase
US-09-907-907A-45
Query Match 24.1%; Score 7; DB 10; Length 245;
Best Local Similarity 100.0%; Pred. No. 58;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 AELEKEV 23
DB 90 AELEKEV 96

RESULT 12
US-10-369-493-3657
; Sequence 3657, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF

US-10-739-930-10206
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 3657
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Neurospora crassa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(284)
; OTHER INFORMATION: unsure at all xaa locations
US-10-369-493-3657
Query Match 24.1%; Score 7; DB 15; Length 284;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 VAELEKE 22
DB 58 VAELEKE 64

RESULT 13
US-10-424-599-250565
; Sequence 250565, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 250565
; LENGTH: 286
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_68290C.1.pep
US-10-424-599-250565
Query Match 24.1%; Score 7; DB 15; Length 286;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 RTEVAEL 19
DB 123 RTEVAEL 129

RESULT 14
US-10-424-599-229602
; Sequence 229602, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
```



```
; SEQ ID NO 229602
; LENGTH: 290
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_49355C.1.pap
US-10-424-599-229602

Query Match      24.1%; Score 7; DB 15; Length 290;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      16 VAELEKE 22
Db      140 VAELEKE 146

RESULT 15
US-10-437-963-108061
; Sequence 108061, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 108061
; LENGTH: 350
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_12351C.1.pap
US-10-437-963-108061

Query Match      24.1%; Score 7; DB 16; Length 350;
Best Local Similarity 100.0%; Pred. No. 80;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 AFLEKEN 9
Db      343 AFLEKEN 349

Search completed: July 26, 2005, 12:28:15
Job time : 69.2708 secs
```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 26, 2005, 12:00:54 ; Search time 13.2917 Seconds
(without alignments)
209.927 Million cell updates/sec

Title: US-10-624-218-1

Perfect score: 29

Sequence: 1 RAAFLERKENTALTREVALEKEVGRCENI 29

Scoring table:

OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 283416 seqs, 96216763 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

PIR 79:*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	19	65.5	261	2 A40579	trans-activating t
2	19	65.5	293	2 S50109	vitellogenin gene-
3	19	65.5	303	2 G02360	thyrotroph embryon
4	19	65.5	313	2 A41524	transcription fact
5	9	31.0	325	2 A55558	albumin D-box bind
6	8	27.6	335	1 DEBSGF	glyceraldehyde-3-p
7	7	24.1	74	2 B82597	hypothetical prote
8	7	24.1	114	2 D72618	hypothetical prote
9	7	24.1	219	2 T45997	hypothetical prote
10	7	24.1	228	2 T38622	ribulose-phosphate
11	7	24.1	234	2 A82814	thiamin-phosphate
12	7	24.1	287	2 AC1473	transcription regu
13	7	24.1	287	2 AG1111	transcription regu
14	7	24.1	320	2 AP2714	conserved hypothet
15	7	24.1	352	2 C97496	hypothetical prote
16	7	24.1	357	2 G81323	translation elonga
17	7	24.1	388	2 AB1371	aminotransferase h
18	7	24.1	636	2 T35182	probable ABC-type
19	7	24.1	845	2 I48176	synaptonemal compl
20	7	24.1	1070	2 T08733	kinasin homolog f2
21	7	24.1	1310	1 I53597	proline dehydrogen
22	7	24.1	1320	1 S66279	proline dehydrogen
23	7	24.1	1320	2 AE0633	proline dehydrogen
24	6	20.7	63	2 E41608	hypothetical prote
25	6	20.7	83	1 B69490	ribosomal protein
26	6	20.7	90	2 S11976	acyl carrier prote
27	6	20.7	92	2 JN0827	acyl carrier prote
28	6	20.7	104	2 AF0450	insertion element
29	6	20.7	105	2 D95108	conserved hypothet

30	6	20.7	110	2 T49383	related to potassi
31	6	20.7	114	2 E97976	conserved hypothet
32	6	20.7	115	2 E75105	hypothetical prote
33	6	20.7	120	2 AE0976	conserved hypothet
34	6	20.7	120	2 S47823	hypothetical 13.7k
35	6	20.7	120	2 F86035	hypothetical prote
36	6	20.7	120	2 F91188	hypothetical prote
37	6	20.7	121	2 H71351	probable ribosomal
38	6	20.7	136	2 AH3394	lactoyglutathione
39	6	20.7	136	2 D71029	hypothetical prote
40	6	20.7	137	2 JC4856	alpha-amylose inh
41	6	20.7	138	2 T49183	hypothetical prote
42	6	20.7	141	2 E95138	galactose-6-phosph
43	6	20.7	141	2 C98006	galactose-6-phosph
44	6	20.7	141	2 B64393	hypothetical prote
45	6	20.7	149	2 JC6202	ribosomal protein

ALIGNMENTS

RESULT 1

A40579

trans-activating transcription regulator TEF - rat

C:Species: Rattus norvegicus (Norway rat)

C>Date: 05-Jun-1992 #sequence_revision 05-Jun-1992 #text_change 09-Jul-2004

C:Accession: A40579

R:Drolet, D.W.; Scully, K.M.; Simmons, D.M.; Wegner, M.; Chu, K.; Swanson, L.W.; Rosenf-

Genes Dev. 5; 1739-1753, 1991

A>Title: TEF, a transcription factor expressed specifically in the anterior pituitary d

A:Reference number: A40579; MUID:92009166; PMID:1916262

A:Accession: A40579

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-261 <DRO>

A:Cross-references: UNIPROT:P41224; GB:S58745; NID:G237084; PIDN:AAB20032.1; PID:G23708

C:Keywords: DNA binding; transcription regulation

Query Match 65.5%; Score 19; DB 2; Length 261;

Best Local Similarity 100.0%; Pred. No. 8.3e-12;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RAAFLERKENTALTREVAEL 19

Db 222 RAAFLERKENTALTREVAEL 240

RESULT 2

S50109

vitellogenin gene-binding protein VBP, beta/beta isoform - chicken

C:Species: Gallus gallus (chicken)

C>Date: 14-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 09-Jul-2004

C:Accession: S50109; S50111

R:Burch, J.B.E.; Davis, D.L.

Nucleic Acids Res. 22, 4733-4741, 1994

A>Title: Alternative promoter usage and splicing options result in the differential exp.

ora.

A:Reference number: S50109; MUID:95075656; PMID:7984425

A:Accession: S50109

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-233 <BUR>

A:Cross-references: UNIPROT:Q92172; EMBL:U09221; NID:G483937; PIDN:AAA82156.1; PID:G483

A:Accession: S50111

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-34 <BU2>

A:Cross-references: EMBL:U09223; NID:G483823; PIDN:AAA82158.1; PID:G483824

C:Keywords: transcription factor

Query Match 65.5%; Score 19; DB 2; Length 293;

Best Local Similarity 100.0%; Pred. No. 9.3e-12;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RAAFLKENTALRTEVAEL 19
|||||
Db 274 RAAFLKENTALRTEVAEL 292

RESULT 5
A55558
albumin D-box binding protein - human
C:Species: Homo sapiens (man)
C:Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 09-Jul-2004
C:Accession: A55558; G02887
R:Khatib, Z.A.; Inaba, T.; Valentine, M.; Look, A.T.
Genomics 23, 344-351, 1994
A:Title: Chromosomal localization and cDNA cloning of the human DBP and TEF genes.
A:Reference number: A55558; MUID:95137580; PMID:7835883
A:Accession: A55558
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-325 <KHA>
A:Cross-references: UNIPROT:Q10586; GB:U06936; NID:G606798; PIDN:AAA81374.1; PID:G606799
R:Mueller, C.R.
submitted to the EMBL Data Library, February 1996
A:Reference number: H01797
A:Accession: G02887
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-178, 'S', 180-325 <MUE>
A:Cross-references: EMBL:U48213; NID:g1208753; PID:g1208754
C:Genetics:
A:Gene: GDB:DBP
A:Cross-references: GDB:128840; OMIM:124097
A:Map position: 4q12-4q12
A:Introns: 47/1; 184/1; 254/3

Query Match 31.0%; Score 9; DB 2; Length 325;
Best Local Similarity 100.0%; Pred. No. 0.12; 0; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches

QY 1 RAAFLKENTALRTEVAEL 9
|||||
Db 286 RAAFLKENTALRTEVAEL 294

RESULT 6
DBSSGF
glyceraldehyde-3-phosphate dehydrogenase (phosphorylating) (EC 1.2.1.12) [validated] - Bac
N:Alternate names: triosephosphate dehydrogenase
C:Species: Bacillus stearothermophilus
C:Date: 24-Apr-1984 #sequence_revision 17-Nov-2000 #text_change 09-Jul-2004
C:Accession: JS0164; PS0343; A93186; A91096; A00374
R:Branlant, C.; Oster, T.; Branlant, G.
Gene 75, 145-155, 1989
A:Title: Nucleotide sequence determination of the DNA region coding for Bacillus stearo
n in Escherichia coli.
A:Reference number: JS0164; MUID:89252911; PMID:2656407
A:Accession: JS0164
A:Molecule type: DNA
A:Residues: 1-335 <BRA>
A:Cross-references: UNIPROT:P00362; GB:M24493; NID:g142951; PIDN:AAA22461.1; PID:g14295
R:Davies, G.J.; Littlechild, J.A.; Watson, H.C.; Hall, L.
Gene 109, 39-45, 1991
A:Title: Sequence and expression of the gene encoding 3-phosphoglycerate kinase from Bac
A:Reference number: JQ1399; MUID:92097950; PMID:1756980
A:Accession: PS0343
A:Molecule type: DNA
A:Residues: 315-335 <DAV>
A:Cross-references: EMBL:X58059; NID:g48853; PIDN:CAA41092.1; PID:g48854
A:Experimental source: strain NCAL503
R:Biesecker, G.; Harris, J.I.; Thierry, J.C.; Walker, J.E.; Wonacott, A.J.
Nature 266, 328-333, 1977
A:Title: Sequence and structure of D-glyceraldehyde 3-phosphate dehydrogenase from Bac
A:Reference number: A93186; MUID:77171226; PMID:193030
A:Contents: sequence; X-ray crystallography, 2.7 angstroms

A:Accession: A93186
A:Molecule type: protein
A:Residues: 2-36, 'N', 38, 'DG', 41-58, 'V', 60-61, 'DGDVS', 67-78, 'N', 80-124, 'V', 126, 'N', 128-132
R:Walker, J.E.; Carne, A.F.; Runswick, M.J.; Bridgen, J.; Harris, J.I.
Eur. J. Biochem. 108, 549-565, 1980
A:Title: D-Glycerinaldehyde-3-phosphate dehydrogenase. Complete amino-acid sequence of the
A:Reference number: A91096; MUID:81003878; PMID:7408868
A:Accession: A91096
A:Molecule type: protein
A:Residues: 2-36, 'N', 38, 'DG', 41-58, 'V', 60-61, 'DGDVS', 67-78, 'N', 80-124, 'V', 126, 'N', 128-132
C:Genetics:
A:Gene: gap
C:Superfamily: glycerinaldehyde-3-phosphate dehydrogenase
C:Keywords: gluconeogenesis; glycolysis; homotetramer; NAD; oxidoreductase
F:4-34/Region: beta-alpha-beta NAD nucleotide-binding fold
F:152/Active site: Cys #status experimental
F:179/Active site: His #status predicted

Query Match 27.6%; Score 8; DB 1; Length 335;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 VAELEKEV 23
|||||
DB 245 VAELEKEV 252

RESULT 7
B82597
hypothetical protein XP2130 [imported] - Xylella fastidiosa (strain 9a5c)
C:Species: Xylella fastidiosa
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C:Accession: B82597
R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; MUID:20365717; PMID:10910347
A:Note: for a complete list of authors see reference number A59328 below
A:Accession: B82597
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-74 <Sim>
A:Cross-references: UNIPROT:Q9PBL4; GB:AE004027; GB:AE003849; NID:99107249; PIDN:AAF8492
A:Experimental source: strain 9a5c
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H
as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
A:Authors: Marting, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.B.; de Sa, R.G.; Santelli, R.V.; Sawasak
A:Authors: da Silva, A.C.R.; da Silva, P.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Teuhako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: XP2130

Query Match 24.1%; Score 7; DB 2; Length 74;
Best Local Similarity 100.0%; Pred. No. 3.1;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 RTEVAEL 19
|||||
DB 17 RTEVAEL 23

RESULT 8
D72618
hypothetical protein APE1405 - Aeropyrum pernix (strain K1)

C:Species: Aeropyrum pernix
C>Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C:Accession: D72618
R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K
DNA Res. 6, 83-101, 1999
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropy
A:Reference number: A72450; MUID:99310339; PMID:10382966
A:Accession: D72618
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-114 <KAW>
A:Cross-references: UNIPROT:Q9YC46; DDBJ:AP000061; NID:95104821; PIDN:BAA80402.1; PID:d1
A:Experimental source: strain K1
C:Genetics:
A:Gene: APE1405

Query Match 24.1%; Score 7; DB 2; Length 114;
Best Local Similarity 100.0%; Pred. No. 4.6;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LEKEVGR 25
|||||
DB 48 LEKEVGR 54

RESULT 9
T45997
hypothetical protein F9D24.280 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 09-Jul-2004
C:Accession: T45997
R:D'Angelo, M.; Vezzi, A.; Modesto, D.; Pigazzi, M.; Valle, G.; Mewes, H.W.; Lemcke, K.
submitted to the Protein Sequence Database, January 2000
A:Reference number: Z23011
A:Accession: T45997
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-219 <DAN>
A:Cross-references: UNIPROT:Q9M2H9; EMBL:AL137081
A:Experimental source: cultivar Columbia; BAC clone F9D24
C:Genetics:
A:Map position: 3
A:introns: 85/3
A:Note: F9D24.280
C:Superfamily: Arabidopsis thaliana hypothetical protein F9D24.210

Query Match 24.1%; Score 7; DB 2; Length 219;
Best Local Similarity 100.0%; Pred. No. 8.6;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 AFLEKEN 9
|||||
DB 197 AFLEKEN 203

RESULT 10
T38622
ribulose-phosphate 3-epimerase - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: T38622
R:McLean, J.; Harris, D.; Barrell, B.G.; Rajandream, M.A.; Wood, V.
submitted to the EMBL Data Library, September 1997
A:Reference number: Z21722
A:Accession: T38622
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-228 <MCL>
A:Cross-references: UNIPROT:O14105; EMBL:Z98979; PIDN:CAB11689.1; GSPDB:GN00066; SPDB:S
A:Experimental source: strain 972h-; cosmid c31G5
C:Genetics:
A:Gene: SPDB:SPAC31G5.05c

A;Residues: 1-287 <GLA>
A;Cross-references: UNIPROT:Q92EY6; GB:AL592022; PIDN:CAC95555.1; PID:g16412751; GSPDB:(C)
A;Experimental source: strain Clip11262
C;Genetics:
A;Gene: lin0322

	Query Match	24.1%	Score 7;	DB 2;	Length 287;
Best Local Similarity	100.0%;	Pred. No. 11;			
Matches	7;	Conservative	0;	Mismatches	0;
Indels			0;	Gaps	0;

Qy 16 VAELEKE 22
| | | | |
Db 142 VAELEKE 148

RESULT 13
AG111
transcription regulator lysR-glr family homolog lmo0294 [imported] - Listeria monocytogenes
C;Species: Listeria monocytogenes
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C;Accession: AG111
P;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Aussurget, O.; Entian, K.D.; Fsihi, H.
. J. Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kretz, J.; Kuhn, M.; Kunat, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Maier, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend, A.; Title: Comparative genomics of Listeria species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AG111
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-287 <GLA>
A;Cross-references: UNIPROT:QBYA65; GB:NC_003210; PIDN:CAD00821.1; PID:g16409659; GSPDB:(C)
A;Experimental source: strain EGD-e
C;Genetics:
A;Gene: lmo0294

	Query Match	24.1%	Score 7;	DB 2;	Length 287;
Best Local Similarity	100.0%;	Pred. No. 11;			
Matches	7;	Conservative	0;	Mismatches	0;
Indels			0;	Gaps	0;

Qy 16 VAELEKE 22
| | | | |
Db 142 VAELEKE 148

RESULT 14
AF2714
conserved hypothetical protein Atull19 [imported] - Agrobacterium tumefaciens (strain C)
C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C;Accession: AF2714
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, T.; Erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClellan, P.; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, S.; Ew
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A;Reference number: AB2577; MUID:21608550; PMID:11743193
A;Accession: AF2714
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-320 <KUR>
A;Cross-references: UNIPROT:Q8UGB9; GB:AE008688; PIDN:AAL42132.1; PID:g17739517; GSPDB:(C)
A;Experimental source: strain C58 (Dupont)
C;Genetics:
A;Gene: Atull19
A;Map position: circular chromosome

	Query Match	24.1%	Score 7;	DB 2;	Length 320;
Best Local Similarity	100.0%;	Pred. No. 12;			

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RAAFLEK 7
|||||
Db 153 RAAFLEK 159

RESULT 15
C97496
hypothetical protein AGR_C_2073 [imported] - Agrobacterium tumefaciens (strain C58, Cere
C:Species: Agrobacterium tumefaciens
C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
C:Accession: C97496
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A:Reference number: A97359; MUID:21608551; PMID:11743194
A:Accession: C97496
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-352 <KUR>
A:Cross-references: UNIPROT:Q8UCB9; GB:AE007869; PIDN:AAK86924.1; PID:GI5156152; GSPDB:G
C:Genetics:
A:Gene: AGR_C_2073
A:Map position: circular chromosome

Query Match 24.1%; Score 7; DB 2; Length 352;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RAAFLEK 7
|||||
Db 185 RAAFLEK 191

Search completed: July 26, 2005, 12:20:48
Job time : 15.2917 secs

This Page Blank (uspio)

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	19	65.5	140	2	Q7Z3J7	Q723J7	homo sapien
2	19	65.5	271	1	TEF_RAT	P41224	rattus norv
3	19	65.5	273	2	Q7Z5X6	Q725X6	homo sapien
4	19	65.5	293	2	Q92172	Q92172	gallus gall
5	19	65.5	301	1	TEF_MOUSE	Q9J1C6	mus musculus
6	19	65.5	301	1	TEF_HUMAN	Q10587	homo sapien
7	17	58.6	221	1	TEF_PHOSU	P97516	phodopus su
8	13	44.8	278	2	Q7S7V6	Q7SV76	xenopus lae
9	13	44.8	298	2	Q632P9	Q632P9	xenopus lae
10	9	31.0	188	2	Q6R2I2	Q6R2I2	rattus norv
11	9	31.0	225	2	Q6R2I3	Q6R2I3	rattus norv
12	9	31.0	325	1	DBP_HUMAN	Q10586	homo sapien
13	9	31.0	325	1	DBP_MOUSE	Q60925	mus musculus
14	9	31.0	325	1	DBP_RAT	P16443	rattus norv
15	8	27.6	334	1	G3P_BACST	P00362	bacillus st
16	8	27.6	534	2	O33898	O33898	streptococc
17	8	27.6	534	2	O68165	O68165	streptococc
18	8	27.6	548	2	Q76EM4	Q76EM4	gluconobact
19	8	27.6	570	2	Q9XDC5	Q9XDC5	streptococc
20	8	27.6	570	2	O8NZ4A	O8NZ4A	streptococc
21	8	27.6	785	2	O6F293	O6F293	mesoplasma
22	8	27.6	1068	1	DAM2_HUMAN	Q86165	homo sapien
23	7	24.1	74	2	Q9PB14	Q9PB14	xyliella fas
24	7	24.1	104	2	O83BI1	O83BI1	coxiella bu
25	7	24.1	114	2	O9YCA6	O9YCA6	aeropyrum p
26	7	24.1	115	2	O6JSU2	O6JSU2	siphonocybe
27	7	24.1	116	2	O6JSX7	O6JSX7	docodeemus
28	7	24.1	116	2	Q9BN95	Q9BN95	zercon sp.
29	7	24.1	116	2	Q9BN98	Q9BN98	tachyuropod
30	7	24.1	116	2	Q9BN8A	Q9BN8A	sejus sp.
31	7	24.1	116	2	Q9BNC3	Q9BNC3	parasitid

SEQUENCE FROM N.A., INTERACTION WITH DBP, AND MUTAGENESIS.
RC TISSUE=Pituitary;
RX MEDLINE=92009166; PubMed=1916262;
RA Drolet D.W., Scully K.M., Simmons D.M., Wegner M., Chu K.,
Swanson L.W., Rosenfeld M.G.;
RT "TEF, a transcription factor expressed specifically in the anterior
pituitary during embryogenesis, defines a new class of leucine zipper
proteins";
RL Genes Dev. 5:1739-1753(1991).
CC -!- FUNCTION: Transcription factor that binds to and transactivates
[TC] [AG] [AG] [TATC] [AG]-3'.
CC -!- SUBUNIT: Binds DNA as a homodimer or a heterodimer. Can form a
heterodimer with DBP.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- TISSUE SPECIFICITY: Expressed exclusively in the rostral portion
of the anterior pituitary during embryogenesis. Found in several
tissues in juvenile and adult rats.
CC -!- DEVELOPMENTAL STAGE: Expressed up to embryonic day 14 and
specifically in the anterior pituitary during embryogenesis.
CC -!- INDUCTION: Accumulates according to a robust circadian rhythm (By
similarity).
CC -!- SIMILARITY: Belongs to the bZIP family. PAR subfamily.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; S58745; AAB20032.1; ALT_INIT.
PIR; A40579; A40579.
DR TRANSEAC; T01072; --
DR RGD; 3841; Tef.
DR InterPro; IPR004827; TF_bZIP.
DR Pfam; PF00170; bZIP; 1.
DR SMART; SM00338; BRLZ; 1.
DR PROSITE; PS50217; bZIP; 1.
DR PROSITE; PS00036; bZIP BASIC; FALSE NEG.
KW Activator; Biological rhythms; DNA-binding; Nuclear protein;
Transcription regulation.
FT NON_TER 1 1
FT DOMAIN 134 185 Pro-rich (Proline/Acidic region (PAR)).
FT DNA_BIND 216 225 Basic motif.
FT DOMAIN 236 250 K->A: 30-fold decrease in affinity for
FT MUTAGEN 188 188 prolactin recognition element.
FT MUTAGEN 189 189 K->A: 30-fold decrease in affinity for
FT MUTAGEN 191 191 prolactin recognition element.
FT MUTAGEN 192 192 K->A: 30-fold decrease in affinity for
FT MUTAGEN 192 192 prolactin recognition element.
FT MUTAGEN 236 236 L->V: Diminishes DNA-binding.
FT MUTAGEN 243 243 L->V: Diminishes DNA-binding.
FT SEQUENCE 271 AA; 30330 MW; 1A8CF2396C1188B9 CRC64;
Query Match 65.5%; Score 19; DB 1; Length 271;
Best Local Similarity 100.0%; Pred. No. 1.4e-10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RAAFLKENTALTREVAEL 19
ID 232 RAAFLKENTALTREVAEL 250
Db
RESULT 3
Q726X6 PRELIMINARY; PRT; 273 AA.
AC Q726X6;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)

DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE DJ979N1.5.2 (Thyrotrophic embryonic factor (Orthlog of chicken
vitellogenin gene-binding protein VBP beta/beta variant) (Variant
2)).
GN Name=TEF;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Smith M.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the bZIP family.
DR EMBL; AL035659; CAB62497.1; --
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR004827; TF_bZIP.
DR SMART; SM00338; BRLZ; 1.
DR PROSITE; PS50217; bZIP; 1.
KW DNA-binding; Nuclear protein.
SQ SEQUENCE 273 AA; 30634 MW; B1A956039C8C4AC4 CRC64;
Query Match 65.5%; Score 19; DB 2; Length 273;
Best Local Similarity 100.0%; Pred. No. 1.4e-10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RAAFLKENTALTREVAEL 19
ID 234 RAAFLKENTALTREVAEL 252
Db
RESULT 4
Q92172 PRELIMINARY; PRT; 293 AA.
AC Q92172;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Vitellogenin binding protein (VBP), beta/beta isoform.
GN Name=Vbp;
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=White Leghorn;
RX MEDLINE=95075656; PubMed=7984425;
RA Burch J.B., Davis D.L.;
RT "Alternative promoter usage and splicing options result in the
differential expression of mRNAs encoding four isoforms of chicken
VBP, a member of the PAR subfamily of bZIP transcription factors.";
RT Nucleic Acids Res. 22:4733-4741(1994).
RL -!- SIMILARITY: Belongs to the bZIP family.
CC EMBL; U09221; AAA82156.1; --
DR PIR; S50109; S50109.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR SMART; SM00338; BRLZ; 1.
DR PROSITE; PS50217; bZIP; 1.
KW DNA-binding; Nuclear protein.
SQ SEQUENCE 293 AA; 32736 MW; C5803D510CEB03D6 CRC64;
Query Match 65.5%; Score 19; DB 2; Length 293;
Best Local Similarity 100.0%; Pred. No. 1.5e-10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RAAFLKENTALTREVAEL 19

DB 248 RAAFLKENTALTREVAEL 266
 RESULT 5
 ID TEF_MOUSE STANDARD; PRT; 301 AA.
 AC Q9JIC6; Q6QHT6; Q8C610; Q8VD02;
 DT 25-OCT-2004 (Rel. 45, Created)
 DT 25-OCT-2004 (Rel. 45, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Thyrotroph embryonic factor.
 GN Name=Tef;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM ALPHA).
 RC STRAIN=BALB/c; TISSUE=Brain;
 RX MEDLINE=20261582; PubMed=10799536; DOI=10.1074/jbc.275.19.14524;
 RA Krueger D.A., Warner E.A., Dowd D.R.;
 RT "Involvement of thyrotroph embryonic factor in calcium-mediated
 regulation of gene expression.";
 RL J. Biol. Chem. 275:14524-14531 (2000).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORMS ALPHA AND BETA), AND ALTERNATIVE PROMOTER
 USAGE.
 RX PubMed=14702338; DOI=10.1074/jbc.M313822200;
 RA Zhou J., Hoggatt A.M., Herring B.P.;
 RT "Activation of the smooth muscle-specific telokin gene by thyrotroph
 embryonic factor (TEF)."
 RL J. Biol. Chem. 275:15929-15937 (2004).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORMS ALPHA AND 2).
 RC STRAIN=FVB/N; TISSUE=Breast tumor, and Kidney;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Klausner R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Strausberg R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.P., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalek U., Smalish D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [4]
 RP SEQUENCE OF 97-196 FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Kidney;
 RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
 RA Nikaido I., Osego N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojohori T.,
 RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
 RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
 RA Blake J.A., Bradt D., Brusci V., Chothia C., Corbani L.E., Cousins S.,
 RA Dalla E., Dragani T.A., Fletcher C.P., Forrest A., Frazer K.S.,
 RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
 RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
 RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.B.,
 RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
 RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
 RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
 RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
 RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
 RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Welle C.,
 RA Wilming L.G., Wyshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
 RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
 RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
 RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
 RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
 RA Birney E., Hayashizaki Y.;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 60,770 full-length cDNAs.";
 RL Nature 420:563-573 (2002).
 RN [5]
 RP CIRCADIEN INDUCTION.
 RX PubMed=8617210;
 RA Fonjallaz P., Ossipow V., Wanner G., Schibler U.;
 RT "The two PAR leucine zipper proteins, TEF and DBP, display similar
 circadian and tissue-specific expression, but have different target
 promoter preferences.";
 RL EMBO J. 15:351-362 (1996).
 RN [6]
 RP IMPLICATION IN EPILEPSY.
 RX PubMed=15175240; DOI=10.1101/gad.301404;
 RA Gachon P., Fonjallaz P., Damiola F., Gos P., Kodama T., Zakany J.,
 RA Duboule D., Petit B., Tafti M., Schibler U.;
 RT "The loss of circadian PAR bzip transcription factors results in
 epilepsy.";
 RL Genes Dev. 18:1397-1412 (2004).
 CC -!- FUNCTION: Transcription factor that binds to and transactivates
 the TSHB promoter. Binds to a minimal DNA-binding sequence 5'-
 [TC] [AG] [AG] TTA [TC] [AG] -3' (By similarity). Also activates the
 telokin promoter in smooth muscle-specific and calcium-dependent
 manner.
 CC -!- SUBUNIT: Binds DNA as a homodimer or a heterodimer. Can form a
 heterodimer with DBP (By similarity).
 CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative promoter;
 CC Comment=3 isoforms, Alpha (shown here), Beta and 2, are produced
 by use of alternative promoters;
 CC -!- TISSUE SPECIFICITY: Isoform Alpha and isoform Beta are expressed
 at high levels in lung, bladder, kidney, gut and brain.
 CC -!- INDUCTION: Accumulates according to a robust circadian rhythm in
 liver and kidney. In liver nuclei, the amplitude of daily
 oscillation has been estimated to be 9-fold. Expressed at nearly
 constant level in the brain.
 CC -!- MISCELLANEOUS: Mice deficient for all three PAR bzip proteins
 (DBP, HLF and TEF) display a dramatically shortened life span and
 are highly susceptible to generalized spontaneous and audiogenic
 epilepsies (due for example to the noise of a vacuum cleaner) that
 are frequently lethal. The down-regulation of pyridoxal kinase
 (pdxk) expression in these mice may participate in this seizure
 phenotype.
 CC -!- SIMILARITY: Belongs to the bZIP family. PAR subfamily.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See http://www.isb-sib.ch/announce/
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF194420; AAF70830.1; -;
 CC EMBL; AY540631; AAS45599.1; -;
 CC EMBL; AY540632; AAS45600.1; -;
 CC EMBL; BC017689; AAH17689.1; -;
 CC EMBL; BC036982; AAH36982.1; -;

Thu Jul 28 11:34:39 2005

```

DR EMBL; AK075601; BAC35849.1; -.
DR MGD; MGI:98663; Tef.
DR GO; GO:0005634; C:nucleus; IC.
DR GO; GO:0003690; F:double-stranded DNA binding; IDA.
DR GO; GO:0005515; F:protein binding; IPI.
DR GO; GO:0016563; F:transcriptional activator activity; IDA.
DR GO; GO:0045944; P:positive regulation of transcription from P. . . ; IDA.
DR InterPro; IPR004827; TF_bZIP.
DR Pfam; PF00170; bZIP; 1.
DR SMART; SM00338; bZIP; 1.
DR PROSITE; PS00217; bZIP; 1.
DR PROSITE; PS00036; bZIP_BASIC; FALSE_NEG.
DR KW Activator; Alternative promoter usage; Biological rhythms;
DNA-binding; Nuclear protein; Transcription regulation.
FT DOMAIN 164 215 pro-rich (Proline/Acidic region (PAR)).
FT DNA_BIND 237 255 Basic motif.
FT DOMAIN 266 280 Leucine-zipper.
FT VARSPLIC 1 50 MSDAGGKKPVEPQAGPGRAAGRLGSGFPLVLKLM
FT ENPRETRL -> MSSCSQIGVAPMDPEVLKSLHSLP
FT WSEKKA (in isoform Beta).
FT /FTID-VSP 011245.
FT MSDAGGKKPVEPQAGPGRAAGRLGSGFPLVLKLM
FT ENPRETRL -> MDMEVLKSLHSLPWSSEKA (in
FT isoform 2).
FT /FTID-VSP 011246.
FT M -> V (in Ref. 4).
FT CONFLICT 97 97
FT SEQUENCE 301 AA; 33145 MW; F83FFFC6D7E091A4 CRC64;
Query Match 65.5%; Score 19; DB 1; Length 301;
Best Local Similarity 100.0%; Pred. No. 1.5e-10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 RAAPLEKENTARTEVAEL 19
Db 262 RAAPLEKENTARTEVAEL 280
|||||
TEF HUMAN STANDARD; PRT; 303 AA.
AC Q10587; Q15729; Q8IU94; Q96TGA;
DT 01-OCT-1996 (Rel. 34, Created)
DT 25-OCT-2004 (Rel. 45, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Thymotroph embryonic factor.
GN Name=TEF;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95137580; PubMed=7835883;
RA Khatib Z.A., Inaba T., Valentine M., Look A.T.;
RT "Chromosomal localization and cDNA cloning of the human DBP and TEF
RL Genomics 23:344-351(1994).
[2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96219638; PubMed=8639829;
RA Hunger S.P., Li S., Fall M.Z., Naumovski L., Cleary M.L.;
RT "The proto-oncogene HLF and the related basic leucine zipper protein
RT TEF display highly similar DNA-binding and transcriptional regulatory
RT properties.";
RL Blood 87:4607-4617(1996).
[3]
RP SEQUENCE FROM N.A.
RA Collins J.E., Wright C.L., Edwards C.A., Davis M.P., Grinham J.A.,
RA Cole C.G., Goward M.E., Aguado B., Mallia M., Mokrab Y., Huckle E.J.,
RA Beare D.W., Dunham I.;
RT "A genome annotation driven approach to cloning the human ORFeome.";
RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.
[4]

```

```

RP SEQUENCE FROM N.A.
RA Halleck A., Ebert L., Moundinya M., Schick M., Eisenstein S.,
RA Neubert P., Kstrang K., Schatten R., Shen B., Henze S., Mar W.,
RA Korn B., Zuo B., Hu Y., Labaer J.;
RT "Cloning of human full open reading frames in Gateway(TM) system entry
RT vector (pDONR201).";
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
[5]
RP SEQUENCE FROM N.A.
RX MEDLINE=20057165; PubMed=10591208; DOI=10.1038/990031;
RA Dunham I., Hunt A.R., Collins J.E., Bruskiewicz R., Beare D.M.,
RA Clamp M., Smink L.J., Ainscough R., Almeida J.P., Babbage A.K.,
RA Bagguley C., Bailey J., Barlow K.F., Bates K.N., Beasley O.P.,
RA Bird C.P., Blakey S.E., Bridgman A.M., Buck D.P., Burgess J.,
RA Burill W.D., Burton J., Carder C., Carter N.P., Chen Y., Clark G.,
RA Clegg S.M., Cobley V.E., Cole C.G., Collier R.E., Connor R.,
RA Conroy D., Corby N.R., Coville G.J., Cox A.V., Davis J., Dawson E.,
RA Dhami P.D., Dockree C., Dodsworth S.J., Durbin R.M., Ellington A.G.,
RA Evans K.L., Fey J.M., Fleming K., French L., Garner A.A., Hall C.,
RA Gilbert J.G.R., Goward M.E., Grafham D.V., Griffiths M.N.D., Hall C.,
RA Hall R.E., Hall-Tamlyn G., Heathcote R.W., Ho S., Holmes S.,
RA Hunt S.E., Jones M.C., Kershaw J., Kimberley A.M., King A.,
RA Laird G.K., Langford C.F., Leversha M.A., Lloyd C., Lloyd D.M.,
RA Martyn I.D., Mashreghi-Mohammadi M., Matthews L.H., Mccann O.T.,
RA Mcclay J., McLaren S., McMurray A.A., Milne S.A., Mortimore B.J.,
RA Phillips S.H., Plumb R.W., Ramsay H., Ramsey Y., Rogers L., Ross M.T.,
RA Scott C.E., Sehra H.K., Skuce C.D., Smalley S., Smith M.L.,
RA Soderlund C., Spragon L., Steward C.A., Sulston J.E., Swann R.M.,
RA Vaudin M., Wall M., Wallis J.M., Whiteley M.N., Willey D.L.,
RA Williams L., Williams S.A., Williamson H., Wilmer T.E., Wilming L.,
RA Wright C.L., Hubbard T., Bentley D.R., Beck S., Rogers J., Shimizu N.,
RA Minoshima S., Kawasaki K., Sasaki T., Asakawa S., Kudoh J.,
RA Shintani A., Shibuya K., Yoshizaki Y., Aoki N., Mitsuyama S.,
RA Roe B.A., Chen F., Chu L., Crabtree J., Deschamps S., Do A., Do T.,
RA Dorman A., Fang F., Fu Y., Hu P., Hua A., Kenton S., Lai H., Lao H.I.,
RA Lewis J., Lewis S., Lin S.-P., Loh P., Malaj E., Nguyen T., Pan H.,
RA Phan S., Qi S., Qian Y., Ray L., Ren Q., Shaull S., Sloan D., Song L.,
RA Wang Q., Wang Y., Wang Z., White J., Willingham D., Wu H., Yao Z.,
RA Zhan M., Zhang G., Chisoe S., Murray J., Miller N., Minx P.,
RA Fulton R., Johnson D., Bemis G., Bentley D., Bradshaw H., Bourne S.,
RA Cordes M., Du Z., Fulton L., Goela D., Graves T., Hawkins J.,
RA Hinds K., Kemp K., Latreille P., Layman D., Ozersky P., Rohlffing T.,
RA Scheet P., Walker C., Wamsley A., Wohldmann P., Pepin K., Nelson J.,
RA Korf I., Bedell J.A., Hillier L.W., Mardis E., Waterston R.,
RA Wilson R., Emanuel B.S., Shaikh T., Kurahashi H., Saitta S.,
RA Budarf M.L., McDermid H.E., Johnson A., Wong A.C.C., Morrow B.E.,
RA Edelman L., Kim U.J., Shizuya H., Simon M.I., Dumanski J.P.,
RA Peyrard M., Kedra D., Serousi E., Fransson I., Tapia I., Bruder C.E.,
RA O'Brien K.P., Wilkinson P., Bodenteich A., Hartman K., Hu X.,
RA Khan A.S., Lane L., Tilahun Y., Wright H.;
RT "The DNA sequence of human chromosome 22.";
RL Nature 402:489-495(1999).
[6]
RP SEQUENCE FROM N.A.
RT TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Berge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.W., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hoeg F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hsieh L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whitling M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Rodriguez A.C., Grimwood J.W., Green E.D., Dickson M.C.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,

```

RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RL and mouse cDNA sequences.";
 CC Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 CC
 CC !- FUNCTION: Transcription factor that binds to and transactivates
 CC the TSHB promoter. Binds to a minimal DNA-binding sequence 5'-
 CC [TC] (AG) [AG] TTA [TC] (AG) -3'. (By similarity).
 CC
 CC !- SUBUNIT: Binds DNA as a homodimer or a heterodimer. Can form a
 CC heterodimer with DBP.
 CC
 CC !- SUBCELLULAR LOCATION: Nuclear.
 CC
 CC !- INDUCTION: Accumulates according to a robust circadian rhythm (By
 CC similarity).
 CC
 CC !- SIMILARITY: Belongs to the bZIP family. PAR subfamily.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; U06935; AA81373.1; ALT_INIT.
 CC EMBL; U40599; AA806497.1; -.
 CC EMBL; CR456592; CAG30478.1; -.
 CC EMBL; CR541827; CAG46626.1; -.
 CC EMBL; AL035659; CAB62498.1; ALT_INIT.
 CC EMBL; BC039258; AAH39258.1; -.
 CC EMBL; BC042476; AAH42476.1; -.
 CC PIR; G02360; G02360.
 CC TRANSFAC; T04876; -.
 CC Genew; HGNC:11722; TEF.
 CC MIM; 188595; -.
 CC GO; GO:0003702; F:RNA polymerase II transcription factor acti. .; TAS.
 CC GO; GO:0006357; P:regulation of transcription from Pol II pro. .; TAS.
 CC InterPro; IPR004827; TF_bZIP.
 CC Pfam; PF00170; bZIP; 1.
 CC SMART; SM00338; BRLZ; 1.
 CC PROSITE; PS0217; bZIP; 1.
 CC PROSITE; PS00036; bZIP_BASIC; FALSE NEG.
 CC Activator; Biological rhythms; DNA-Binding; Nuclear protein;
 KW Transcription regulation.
 FT DOMAIN 166 217 Pro-rich (Proline/Acidic region (PAR)).
 FT DNA BIND 239 257 Basic motif.
 FT DOMAIN 268 282 Leucine-zipper.
 FT CONFLICT 54 54 K -> E (in Ref. 1).
 FT SEQUENCE 303 AA; 33247 MW; 4A87B7BFA7248C6F CRC64;
 SQ
 Query Match 65.5%; Score 19; DB 1; Length 303;
 Best Local Similarity 100.0%; Pred. No. 1.5e-10;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 RAAFLKENTALRTEVAEL 19
 DB 264 RAAFLKENTALRTEVAEL 282
 |||||
 RESULT 7
 TEF PHOSU STANDARD; PRT; 221 AA.
 AC P97516;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Tef-prov protein.
 GN Names:TEF;
 OS Phodopus sungorus (Striped hairy-footed hamster) (Djungarian hamster).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
 OC Phodopus.
 OX NCBI_TaxID=10044;
 RN [1]
 RP SEQUENCE FROM N.A.

RA Bockmann J.;
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
 CC !- FUNCTION: Transcription factor that binds to and transactivates
 CC the TSHB promoter. Binds to a minimal DNA-binding sequence 5'-
 CC [TC] (AG) [AG] TTA [TC] (AG) -3'. (By similarity).
 CC
 CC !- SUBUNIT: Binds DNA as a homodimer or a heterodimer. Can form a
 CC heterodimer with DBP (By similarity).
 CC
 CC !- SUBCELLULAR LOCATION: Nuclear (By similarity).
 CC
 CC !- INDUCTION: Accumulates according to a robust circadian rhythm.
 CC
 CC !- SIMILARITY: Belongs to the bZIP family. PAR subfamily.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; Y11149; CAA72036.1; -.
 CC InterPro; IPR004827; TF_bZIP.
 CC SMART; SM00338; BRLZ; 1.
 CC PROSITE; PS0217; bZIP; 1.
 CC PROSITE; PS00036; bZIP_BASIC; FALSE NEG.
 CC Activator; Biological rhythms; DNA-Binding; Nuclear protein;
 KW Transcription regulation.
 FT NON_TER 1 1
 FT DOMAIN 106 157 Pro-rich (Proline/Acidic region (PAR)).
 FT DNA BIND 179 197 Basic motif.
 FT DOMAIN 208 >221 Leucine-zipper.
 FT NON_TER 221 221
 FT SEQUENCE 221 AA; 24528 MW; 69645BCA042CFA8B CRC64;
 SQ
 Query Match 58.6%; Score 17; DB 1; Length 221;
 Best Local Similarity 100.0%; Pred. No. 1.1e-08;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 RAAFLKENTALRTEVA 17
 DB 204 RAAFLKENTALRTEVA 220
 |||||
 RESULT 8
 Q7SY76 PRELIMINARY; PRT; 278 AA.
 AC Q7SY76;
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-WAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Tef-prov protein.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Whole;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.W., Schuler G.D.,
 RA Hochschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McSwain P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

```

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RA "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Whole;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RA "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Whole;
RX Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the bZIP family.
DR EMBL; BC054981; AAH54981.1; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR004827; TF_bZIP.
DR SMART; SM00338; BRLZ; 1.
DR PROSITE; PS0217; BZIP; 1.
KW DNA-binding; Nuclear protein.
SQ SEQUENCE 278 AA; 31536 MW; 7DA75119B3F24509 CRC64;

Query Match 44.8%; Score 13; DB 2; Length 278;
Best Local Similarity 100.0%; Pred. No. 0.00011;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RAAFLKENTALR 13
DB 239 RAAFLKENTALR 251

RESULT 9
ID Q632P9 PRELIMINARY; PRT; 298 AA.
AC Q632P9;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RA "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heieh F.,
RA Diatchenko L., Narusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Butterfield Y.S.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Schnerch A., Schein J.E.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RA "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC082861; AAH82861.1; -.
KW Hypothetical protein.
SQ SEQUENCE 298 AA; 33649 MW; D544B58FCEA882D4 CRC64;

Query Match 44.8%; Score 13; DB 2; Length 298;
Best Local Similarity 100.0%; Pred. No. 0.00012;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RAAFLKENTALR 13
DB 253 RAAFLKENTALR 265

RESULT 10
ID Q6R2I2 PRELIMINARY; PRT; 188 AA.
AC Q6R2I2;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE D-binding protein 3.
GN Name=Dbb;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Wistar; TISSUE=Brain;
RA Klugmann M., Leichlehn C.B., During M.J.;
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the bZIP family.
DR EMBL; AY518349; AAR99622.1; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR004827; TF_bZIP.
DR SMART; SM00338; BRLZ; 1.
DR PROSITE; PS0217; BZIP; 1.
KW DNA-binding; Nuclear protein.
SQ SEQUENCE 188 AA; 21040 MW; EBCED3EAC7E59478 CRC64;

Query Match 31.0%; Score 9; DB 2; Length 188;
Best Local Similarity 100.0%; Pred. No. 0.65;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RAAFLKENT 9
DB 149 RAAFLKENT 157

RESULT 11
Q6R2I3 PRELIMINARY; PRT; 225 AA.
ID Q6R2I3
AC Q6R2I3;

```

DT 05-JUL-2004 (TReMBLrel. 27, Created)
 DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
 DE D-binding protein 2.
 GN Name=Dbp;
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Wislar; TISSUE=Brain;
 RA Klugmann M., Leitchlein C.B., During M.J.;
 RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: Belongs to the bZIP family.
 DR EMBL; AV518348; AAR99621.1; -;
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0003677; F:DNA binding; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro; IPR004827; TF_bZIP.
 DR SMART; SM00338; BRLZ; 1.
 DR PROSITE; PS50217; BZIP; 1.
 KW DNA-binding; Nuclear protein.
 SQ SEQUENCE 225 AA; 24605 MW; 53378EB8DA0B903E CRC64;
 Query Match 31.0%; Score 9; DB 2; Length 225;
 Best Local Similarity 100.0%; Pred. No. 0.76;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 RAAFLKEN 9
 Db 186 RAAFLKEN 194
 RESULT 12
 ID DBP HUMAN STANDARD; PRT; 325 AA.
 AC Q10586;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE D-site-binding protein (Albumin D box-binding protein) (TAXREB302).
 GN Name=DBP;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95137580; PubMed=7835883;
 RA Khatib Z.A., Inaba T., Valentine M., Look A.T.;
 RT "Chromosomal localization and cDNA cloning of the human DBP and TEF
 genes.";
 RL Genomics 23:344-351(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96374825; PubMed=87861133; DOI=10.1006/geno.1996.0295;
 RA Shuler G., Glasco T., Kang X., Korneluk R., Mueller C.R.;
 RT "Genomic structure of the human D-site binding protein (DBP) gene.";
 RL Genomics 34:334-339(1996).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=97264341; PubMed=9110174;
 RA Yu W., Andersson B., Worley K.C., Muzny D.M., Ding Y., Liu W.,
 RA Ricafrente J.Y., Wentland M.A., Lennon G., Gibbs R.A.;
 RT "Large-scale concatenation cDNA sequencing.";
 RL Genome Res. 7:353-358(1997).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Ovary;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toehiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [5]
 RP SEQUENCE OF 132-325 FROM N.A.
 RX MEDLINE=93246252; PubMed=8482542; DOI=10.1016/0378-1119(93)90375-D;
 RA Nyunoya H., Morita T., Sato T., Honma S., Teujimoto A., Shimotohno K.;
 RT "Cloning of a cDNA encoding a DNA-binding protein TAXREB302 that is
 specific for the tax-responsive enhancer of HTLV-I.";
 RL Gene 126:251-255(1993).
 RN [6]
 RP REVISIONS.
 RA Nyunoya H.;
 RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
 RN [7]
 RP REVIEW.
 RX MEDLINE=99439627; PubMed=10508692; DOI=10.1016/S0959-437X(99)00009-X;
 RA Brown S.A., Schibler U.;
 RT "The ins and outs of circadian timekeeping.";
 RL Curr. Opin. Genet. Dev. 9:588-594(1999).
 CC -!- FUNCTION: This transcriptional activator recognizes and binds to
 the sequence 5'-RTTAYGTAAY-3', found in the promoter of genes such
 as albumin, CYP2A4 and CYP2A5. It is not essential for circadian
 rhythm generation, but modulates important clock output genes. May
 be a direct target for regulation by the circadian pacemaker
 component clock. May affect circadian period and sleep regulation.
 CC -!- SUBUNIT: Binds DNA as a homodimer or a heterodimer. Can form a
 heterodimer with TEF.
 CC -!- SUBCELLULAR LOCATION: Nuclear.
 CC -!- TISSUE SPECIFICITY: Ubiquitously expressed. Expressed in the
 suprachiasmatic nuclei (SCN) and in most peripheral tissues, with
 a strong circadian rhythmicity.
 CC -!- SIMILARITY: Belongs to the bZIP family. PAR subfamily.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U06936; AAA81374.1; -;
 CC EMBL; U48213; AAB18668.1; -;
 CC EMBL; U48212; AAB18668.1; JOINED.
 CC EMBL; U79283; AAB50219.1; -;
 CC EMBL; BC011965; AAB11965.1; -;
 CC EMBL; D28468; BAA05833.1; -;
 CC EIR; A55558; A55558.
 CC TRANSFAC; T04875; -;
 CC Genew; HGNC:2697; DBP.
 CC H-InvDB; HIX0015298; -;
 CC MIM; 124097; -;
 CC GO; GO:0003702; F:RNA polymerase II transcription factor acti. .; TAS.
 CC GO; GO:0006357; P:regulation of transcription from Pol II pro. .; TAS.
 CC InterPro; IPR004827; TF_bZIP.
 CC SMART; SM00338; BRLZ; 1.

PubMed=15175240; DOI=10.1101/gad.301404;
Cachon P., Fonjallaz P., Damiola F., Gos P., Kodama T., Zakany J.,
Duboule D., Petit B., Tafti M., Schibler U.;
"The loss of circadian PAR bzip transcription factors results in
epilepsy.";
Genes Dev. 18:1397-1412(2004).
-!- FUNCTION: This transcriptional activator recognizes and binds to
the sequence 5'-RTTAGTAAV-3', found in the promoter of genes such
as albumin, CYP2A4 and CYP2A5. It is not essential for circadian
rhythm generation, but modulates important clock output genes. May
be a direct target for regulation by the circadian pacemaker
component clock. May affect circadian period and sleep regulation
(By similarity).
-!- SUBUNIT: Binds DNA as a homodimer or a heterodimer. Can form a
heterodimer with TEF (By similarity).
-!- SUBCELLULAR LOCATION: Nuclear.
-!- TISSUE SPECIFICITY: Expressed in the suprachiasmatic nuclei (SCN)
and in most peripheral tissues, with a strong circadian
rhythmicity.
-!- INDUCTION: Accumulates according to a robust circadian rhythm in
liver and kidney. In liver nuclei, the amplitude of daily
oscillation has been estimated to be >50-fold, and 2-fold in the
brain.
-!- MISCELLANEOUS: Mice deficient for all three PAR bzip proteins
(DBP, HLF and TEF) display a dramatically shortened life span and
are highly susceptible to generalized spontaneous and audiogenic
epilepsies (due for example to the noise of a vacuum cleaner) that
are frequently lethal. The down-regulation of pyridoxal kinase
(pdxk) expression in these mice may participate in this seizure
phenotype.
-!- SIMILARITY: Belongs to the bZIP family. PAR subfamily.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; U29762; AAA73924.1; -;
EMBL; BC018323; AAH18323.1; -;
MGD; MGI:94866; Dbp.
InterPro; IPR004827; TF_bZIP.
SMART; SM00338; BRLZ; 1.
PROSITE; PS0217; bZIP; 1.
PROSITE; PS0036; bZIP_BASIC; FALSE NEG.
Activator; Biological Rhythms; DNA-Binding; Nuclear protein;
Transcription regulation.
DOMAIN 129 135 Poly-Pro.
DOMAIN 156 161 Poly-Ser. (Proline/Acidic region (PAR)).
DOMAIN 188 239 Pro-rich.
DOMAIN 258 277 Basic motif.
DNA BIND 290 304 Leucine-zipper.
DOMAIN 60 60 S -> T (in Ref. 1).
FT CONFLICT 109 109 L -> F (in Ref. 1).
SQ SEQUENCE 325 AA; 34379 MW; B2B2A3E091845A16 CRC64;
Query Match 31.0%; Score 9; DB 1; Length 325;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RAAFLEKEN 9
Db 286 RAAFLEKEN 294

RESULT 14
DBP_RAT
ID DBP_RAT STANDARD; PRT; 325 AA.
AC P16443;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT

PROSITE; PS0217; bZIP; 1.
DR PROSITE; PS0036; bZIP_BASIC; FALSE NEG.
KW Activator; Biological Rhythms; DNA-Binding; Nuclear protein;
FT Transcription regulation.
FT DOMAIN 129 135 Poly-Pro.
FT DOMAIN 156 161 Pro-rich (proline/acidic region (PAR)).
FT DNA BIND 258 277 Basic motif.
FT DOMAIN 290 304 Leucine-zipper.
FT CONFLICT 179 179 S -> T (in Ref. 1).
FT CONFLICT 245 245 R -> K (in Ref. 5).
SQ SEQUENCE 325 AA; 34349 MW; A6933CE21399ECF3 CRC64;
Query Match 31.0%; Score 9; DB 1; Length 325;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RAAFLEKEN 9
Db 286 RAAFLEKEN 294

RESULT 13
DBP_MOUSE
ID DBP_MOUSE STANDARD; PRT; 325 AA.
AC Q60925; Q8VCX3;
DT 15-JUL-1998 (Rel. 36, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE D-site-binding protein (Albumin D box-binding protein).
GN Name=Dbp;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SV; TISSUE=Liver;
RA Lee Y.H., Oguchi H., Gonzalez F.J.;
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg K.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Faxner A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uddin T.B., Toehiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Pahay J., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RP "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RN CIRCADIAN INDUCTION.
RX PubMed=8617210;
RA Fonjallaz P., Ossipow V., Wanner G., Schibler U.;
FT "The two PAR leucine zipper proteins, TEF and DBP, display similar
RT circadian and tissue-specific expression, but have different target
RT promoter preferences.";
RL EMOB J. 15:351-362(1996).
RN [4]
RP IMPLICATION IN EPILEPSY.

DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE D-site-binding protein (Albumin D box-binding protein) (D site albumin
DE promoter binding protein 1).
GN Name=Dbp;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Wistar; TISSUE=Liver;
RX MEDLINE=90235277; PubMed=2331750;
RA Mueller C.R., Maire P., Schibler U.;
RT "DBP, a liver-enriched transcriptional activator, is expressed late in
RT ontogeny and its tissue specificity is determined
RT posttranscriptionally.";
RL Cell 61:279-291 (1990).
RN [2]
RP REVISIONS.
RA Mueller C.R., Maire P., Schibler U.;
RL Cell 65:915-915 (1991).
RN [3]
RP REVIEW.
RX MEDLINE=99439627; PubMed=10508692; DOI=10.1016/S0959-437X(99)00009-X;
RA Brown S.A., Schibler U.;
RT "The ins and outs of circadian timekeeping.";
RL Curr. Opin. Genet. Dev. 9:588-594 (1999).
CC -!- FUNCTION: This transcriptional activator recognizes and binds to
CC the sequence 5'-RTTAGTAAY-3', found in the promoter of genes such
CC as albumin, CYP2A4 and CYP2A5. It is not essential for circadian
CC rhythm generation, but modulates important clock output genes. May
CC be a direct target for regulation by the circadian pacemaker
CC component clock. May affect circadian period and sleep regulation
CC (By similarity).
CC -!- SUBUNIT: Binds DNA as a homodimer or a heterodimer. Can form a
CC heterodimer with TEF.
CC -!- CELLULAR LOCATION: Nuclear.
CC -!- TISSUE SPECIFICITY: Expressed in the suprachiasmatic nuclei (SCN)
CC and in most peripheral tissues, with a strong circadian
CC rhythmicity.
CC -!- DEVELOPMENTAL STAGE: Expressed late in ontogeny.
CC -!- SIMILARITY: Belongs to the bZIP family. PAR subfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, J03179; AAA41083.1; --
DR F01, A34894; A34894.
DR TRANSFAC; T00183; --
DR RGD; 2491; Dbp.
DR InterPro; IPR004827; TF bZIP.
DR SMART; SM00338; BRLZ; 1.
DR PROSITE; PS50217; BZIP; 1.
DR PROSITE; PS00336; BZIP_BASIC; FALSE NEG.
KW Activator; Biological rhythms; DNA-binding; Nuclear protein;
KW Transcription regulation.
FT DOMAIN 129 135 Poly-Pro.
FT DOMAIN 156 161 Poly-Ser.
FT DOMAIN 188 239 Pro-rich (Proline/Acidic region (PAR)).
FT DNA_BIND 258 277 Basic motif.
FT DOMAIN 290 304 Leucine-zipper.
SQ SEQUENCE 325 AA; 34436 MW; D9B2A53FF18455B2 CRC64;
Query Match 31.0%; Score 9; DB 1; Length 325;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RAAFLKEN 9

Db 286 RAAFLKEN 294
|||||||
RESULT 15
G3P_BACST STANDARD; PRT; 334 AA.
AC P00362;
DT 21-JUL-1986 (Rel. 01, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12) (GAPDH).
DS Name=gap;
GN Bacillus stearothermophilus.
OS Bacteria; Firmicutes; Bacillales; Bacillaceae; Geobacillus.
OX NCBI_TaxID=1422;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90060813; PubMed=2684782; DOI=10.1016/0378-1119(89)90049-8;
RA Testay H.S., Amelunxen R.E., Goldberg I.D.;
RT "Nucleotide sequences of genes encoding heat-stable and heat-labile
RT glyceraldehyde-3-phosphate dehydrogenases; amino acid sequence and
RT protein thermostability.";
RL Gene 82:237-248 (1989).
RN [2]
RP ERRATUM (RETRACTION).
RX MEDLINE=91033059; PubMed=2227448; DOI=10.1016/0378-1119(90)90484-9;
RA Testay H.S., Amelunxen R.E., Goldberg I.D.;
RL Gene 94:144-144 (1990).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=89252911; PubMed=2656407; DOI=10.1016/0378-1119(89)90391-0;
RA Branlant C., Oster T., Branlant G.;
RT "Nucleotide sequence determination of the DNA region coding for
RT Bacillus stearothermophilus glyceraldehyde-3-phosphate dehydrogenase
RT and of the flanking DNA regions required for its expression in
RT Escherichia coli.";
RL Gene 75:145-155 (1989).
RN [4]
RP SEQUENCE.
RX MEDLINE=81003878; PubMed=7408868;
RA Walker J.E., Carne A.F., Runswick M.J., Bridgen J., Harris J.I.;
RT "D-glyceraldehyde-3-phosphate dehydrogenase. Complete amino-acid
RT sequence of the enzyme from Bacillus stearothermophilus.";
RN [5]
RX X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS).
RX MEDLINE=77171226; PubMed=193030;
RA Biesecker G., Harris J.I., Thierry J.C., Walker J.E., Wonacott A.J.;
RT "Sequence and structure of D-glyceraldehyde 3-phosphate dehydrogenase
RT from Bacillus stearothermophilus";
RN [6]
RX X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS).
RX MEDLINE=87226185; PubMed=3586018;
RA Skarzynski T., Moody P.C.E., Wonacott A.J.;
RT "Structure of holo-glyceraldehyde-3-phosphate dehydrogenase from
RT Bacillus stearothermophilus at 1.8-A resolution.";
RN [7]
RX J. Mol. Biol. 193:171-187 (1987).
RP X-RAY CRYSTALLOGRAPHY (2.45 ANGSTROMS).
RX MEDLINE=97318921; PubMed=9175858; DOI=10.1006/jmbi.1997.0998;
RA Didierjean C., Rahuel-Clermont S., Vitoux B., Dideberg O.,
RA Branlant G., Aubry A.;
RT "A crystallographic comparison between mutated glyceraldehyde-3-
RT phosphate dehydrogenases from Bacillus stearothermophilus complexed
RT with either NAD+ or NADP+.";
RN J. Mol. Biol. 268:739-759 (1997).
CC -!- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +
CC NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
CC -!- PATHWAY: Second phase of glycolysis; first step.
CC -!- SUBUNIT: Homotetramer.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.

CC CC -1- SIMILARITY: Belongs to the glyceraldehyde-3-phosphate
CC CC dehydrogenase family.
CC CC -1- CAUTION: Ref.1 sequence was incorrect and retracted in Ref.2.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC CC or send an email to license@isb-sib.ch).
CC CC -----
CC CC EMBL; M24493; AAA22461.1; -.
DR DR PIR; JS0164; DEBSGF.
DR DR PDB; 1DBV; X-ray; O/P/Q/R=1-334.
DR DR PDB; 1GDI; X-ray; O/P/Q/R=1-334.
DR DR PDB; 1NPT; X-ray; O/P/Q/R=1-334.
DR DR PDB; 1NQS; X-ray; A/C/O/Q=1-334.
DR DR PDB; 1NQO; X-ray; O/P/Q/R=1-334.
DR DR PDB; 1NQO; X-ray; A/C/O/Q=1-334.
DR DR PDB; 2DBV; X-ray; O/P/Q/R=1-334.
DR DR PDB; 2GDI; X-ray; O/P/Q/R=1-334.
DR DR PDB; 3DBV; X-ray; O/P/Q/R=1-334.
DR DR PDB; 4DBV; X-ray; O/P/Q/R=1-334.
DR DR InterPro; IPR000173; GAP_dhhdhrogenase.
DR DR InterPro; IPR006424; GAPDH-I.
DR DR Pfam; PF02800; Gp dh C; 1.
DR DR Pfam; PF00044; Gp dh N; 1.
DR DR PRINTS; PR00078; G3PDHHDHGNASE.
DR DR ProDom; PD007761; GAPDH like; 1.
DR DR TIGRFAMs; TIGR01534; GAPDH-I; 1.
DR DR PROSITE; PS00071; GAPDH; 1.
KW 3D-structure; Direct protein sequencing; Glycolysis; NAD;
KW Oxidoreductase.
FT INIT MET 0
FT BINDING 151
FT ACT_SITE 178
FT STRAND 2
FT HELIX 11
FT TURN 21
FT STRAND 26
FT HELIX 37
FT STRAND 46
FT TURN 48
FT STRAND 51
FT STRAND 57
FT TURN 61
FT STRAND 63
FT TURN 67
FT STRAND 69
FT HELIX 79
FT HELIX 84
FT TURN 88
FT STRAND 91
FT STRAND 101
FT HELIX 102
FT TURN 106
FT TURN 107
FT TURN 111
FT STRAND 115
FT STRAND 123
FT STRAND 127
FT TURN 131
FT HELIX 135
FT TURN 140
FT STRAND 145
FT HELIX 151
FT STRAND 169
FT TURN 182
FT STRAND 185
FT TURN 194
FT STRAND 199
FT TURN 200

FT STRAND 208
FT TURN 211
FT HELIX 212
FT TURN 217
FT TURN 218
FT TURN 220
FT TURN 222
FT TURN 223
FT STRAND 225
FT STRAND 233
FT STRAND 239
FT HELIX 247
FT TURN 253
FT STRAND 265
FT TURN 266
FT STRAND 268
FT TURN 269
FT TURN 270
FT STRAND 271
FT STRAND 272
FT HELIX 275
FT TURN 281
FT STRAND 284
FT TURN 285
FT STRAND 286
FT HELIX 291
FT STRAND 294
FT STRAND 295
FT STRAND 299
FT TURN 301
FT STRAND 302
FT STRAND 304
FT TURN 312
FT TURN 314
FT HELIX 315
FT TURN 316
FT TURN 331
FT SEQUENCE 332
SQ SEQUENCE 334 AA; 35944 MW; 7E8D539E25233A2B CRC64;

Query Match 27.6%; Score 8; DB 1; Length 334;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 VAELEKEV 23
| | | | | | | |
Db 244 VAELEKEV 251

Search completed: July 26, 2005, 12:19:35
Job time : 67.6528 secs

Query Match:	65.52%	Indels:	0
DB:	6	Gaps:	0
US-10-624-218-1 (1-29) x AX305813 (1-446)			
Qy	1 ArgAlaAlaPheLeuGluLysGluAsnThrAlaLeuArgThrGluValAlaGluLeu 19		
Db	88 CGGCAGCGTCTCTGGAGAGGAGACACACGCCCTCGGACGGAGGTTCGCGAGCTT 144		
RESULT 2			
S58745	817 bp mRNA linear ROD 07-MAY-1993		
LOCUS	thyrotroph embryonic factor-leucine zipper transcription factor		
DEFINITION	[rats, pituitary, mRNA, 817 nt].		
ACCESSION	S58745		
VERSION	S58745.1		
KEYWORDS	GI:237084		
SOURCE	Rattus sp.		
ORGANISM	Rattus sp.		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.		
AUTHORS	1 (bases 1 to 817)		
TITLE	Drolet,D.W., Scully,K.M., Simmons,D.M., Wegner,M., Chu,K.T., Swanson,L.W. and Rosenfeld,M.G.		
JOURNAL	TEF, a transcription factor expressed specifically in the anterior pituitary during embryogenesis, defines a new class of leucine zipper proteins		
MEDLINE	Genes Dev. 5 (10), 1739-1753 (1991)		
PUBMED	92009156		
REMARK	GenBank staff at the National Library of Medicine created this entry [NCBI gibseq 58745] from the original journal article.		
FEATURES	Location/Qualifiers		
source	1..817		
gene	/organism="Rattus sp."		
CDS	/mol_type="mRNA"		
	/db_xref="taxon:10118"		
	1..817		
	/gene="thyrotroph embryonic factor, TEF"		
	32..817		
	/gene="thyrotroph embryonic factor, TEF"		
	/note="leucine zipper transcription factor; TEF"		
	/codon_start=1		
	/product="thyrotroph embryonic factor"		
	/protein_id="AAB20032.1"		
	/db_xref="GI:237085"		
	/translation="MENPPRETRLDKEKKEKLEDESAASTWVASLMPPIWDKT IPYDGSEFLEYMDLDFLENGIPASPTHLAQNLLLPVALEKESASSTASPPSS STAIQPSFTVSTESSLEKERETSPIDPNCVDFVNFNPDADLVLSVPGGELFN PRKHFAEDLDKPPQMIKKAKKVFVDPQDEKDYWTRKKNNVAAKRSRDARLKENQ ITIRAAFLKENTALRTEVAELRKEVGKCKTIIVSKYETKGYL"		
ORIGIN			
Alignment Scores:	1.15e-11	Length:	817
Pred. No.:	19.00	Matches:	19
Score:	100.00%	Conservative:	0
Percent Similarity:	100.00%	Mismatches:	0
Best Local Similarity:	100.00%	Indels:	0
Query Match:	65.52%	Gaps:	0
DB:	10		
US-10-624-218-1 (1-29) x S58745 (1-817)			
Qy	1 ArgAlaAlaPheLeuGluLysGluAsnThrAlaLeuArgThrGluValAlaGluLeu 19		
Db	695 CGGCAGCATTTCTGGAGAGGAGACACACGCCCTCGGACGGAGGTTCGAGCTT 751		
RESULT 3			
HSU06935	855 bp mRNA linear PRI 16-NOV-1995		
LOCUS	Human thyrotroph embryonic factor (TEF) mRNA, complete cds.		
DEFINITION	U06935		
ACCESSION			
US-10-624-218-1 (1-29) x HSU06935 (1-855)			
Qy	1 ArgAlaAlaPheLeuGluLysGluAsnThrAlaLeuArgThrGluValAlaGluLeu 19		
Db	714 CGGCAGCGCTTCTCTGGAGAGGAGACACACGCCCTCGGACGGAGGTTCGCGAGCTA 770		
ALIGNMENT			
Pred. No.:	1.2e-11	Length:	855
Score:	19.00	Matches:	19
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	65.52%	Indels:	0
DB:	9	Gaps:	0
US-10-624-218-1 (1-29) x HSU06935 (1-855)			
Qy	1 ArgAlaAlaPheLeuGluLysGluAsnThrAlaLeuArgThrGluValAlaGluLeu 19		
Db	714 CGGCAGCGCTTCTCTGGAGAGGAGACACACGCCCTCGGACGGAGGTTCGCGAGCTA 770		
ALIGNMENT			
Pred. No.:	1.2e-11	Length:	855
Score:	19.00	Matches:	19
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	65.52%	Indels:	0
DB:	9	Gaps:	0
US-10-624-218-1 (1-29) x HSU06935 (1-855)			
Qy	1 ArgAlaAlaPheLeuGluLysGluAsnThrAlaLeuArgThrGluValAlaGluLeu 19		
Db	714 CGGCAGCGCTTCTCTGGAGAGGAGACACACGCCCTCGGACGGAGGTTCGCGAGCTA 770		
ALIGNMENT			
Pred. No.:	1.2e-11	Length:	855
Score:	19.00	Matches:	19
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	65.52%	Indels:	0
DB:	9	Gaps:	0
US-10-624-218-1 (1-29) x HSU06935 (1-855)			
Qy	1 ArgAlaAlaPheLeuGluLysGluAsnThrAlaLeuArgThrGluValAlaGluLeu 19		
Db	714 CGGCAGCGCTTCTCTGGAGAGGAGACACACGCCCTCGGACGGAGGTTCGCGAGCTA 770		
ALIGNMENT			
Pred. No.:	1.2e-11	Length:	855
Score:	19.00	Matches:	19
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	65.52%	Indels:	0
DB:	9	Gaps:	0
US-10-624-218-1 (1-29) x HSU06935 (1-855)			
Qy	1 ArgAlaAlaPheLeuGluLysGluAsnThrAlaLeuArgThrGluValAlaGluLeu 19		
Db	714 CGGCAGCGCTTCTCTGGAGAGGAGACACACGCCCTCGGACGGAGGTTCGCGAGCTA 770		
ALIGNMENT			
Pred. No.:	1.2e-11	Length:	855
Score:	19.00	Matches:	19
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	65.52%	Indels:	0
DB:	9	Gaps:	0
US-10-624-218-1 (1-29) x HSU06935 (1-855)			
Qy	1 ArgAlaAlaPheLeuGluLysGluAsnThrAlaLeuArgThrGluValAlaGluLeu 19		
Db	714 CGGCAGCGCTTCTCTGGAGAGGAGACACACGCCCTCGGACGGAGGTTCGCGAGCTA 770		
ALIGNMENT			
Pred. No.:	1.2e-11	Length:	855
Score:	19.00	Matches:	19
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	65.52%	Indels:	0
DB:	9	Gaps:	0
US-10-624-218-1 (1-29) x HSU06935 (1-855)			
Qy	1 ArgAlaAlaPheLeuGluLysGluAsnThrAlaLeuArgThrGluValAlaGluLeu 19		
Db	714 CGGCAGCGCTTCTCTGGAGAGGAGACACACGCCCTCGGACGGAGGTTCGCGAGCTA 770		
ALIGNMENT			
Pred. No.:	1.2e-11	Length:	855
Score:	19.00	Matches:	19
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	65.52%	Indels:	0
DB:	9	Gaps:	0
US-10-624-218-1 (1-29) x HSU06935 (1-855)			
Qy	1 ArgAlaAlaPheLeuGluLysGluAsnThrAlaLeuArgThrGluValAlaGluLeu 19		
Db	714 CGGCAGCGCTTCTCTGGAGAGGAGACACACGCCCTCGGACGGAGGTTCGCGAGCTA 770		
ALIGNMENT			
Pred. No.:	1.2e-11	Length:	855
Score:	19.00	Matches:	19
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	65.52%	Indels:	0
DB:	9	Gaps:	0
US-10-624-218-1 (1-29) x HSU06935 (1-855)			
Qy	1 ArgAlaAlaPheLeuGluLysGluAsnThrAlaLeuArgThrGluValAlaGluLeu 19		
Db	714 CGGCAGCGCTTCTCTGGAGAGGAGACACACGCCCTCGGACGGAGGTTCGCGAGCTA 770		
ALIGNMENT			
Pred. No.:	1.2e-11	Length:	855
Score:	19.00	Matches:	19
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	65.52%	Indels:	0
DB:	9	Gaps:	0
US-10-624-218-1 (1-29) x HSU06935 (1-855)			
Qy	1 ArgAlaAlaPheLeuGluLysGluAsnThrAlaLeuArgThrGluValAlaGluLeu 19		
Db	714 CGGCAGCGCTTCTCTGGAGAGGAGACACACGCCCTCGGACGGAGGTTCGCGAGCTA 770		
ALIGNMENT			
Pred. No.:	1.2e-11	Length:	855
Score:	19.00	Matches:	19
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	65.52%	Indels:	0
DB:	9	Gaps:	0
US-10-624-218-1 (1-29) x HSU06935 (1-855)			
Qy	1 ArgAlaAlaPheLeuGluLysGluAsnThrAlaLeuArgThrGluValAlaGluLeu 19		
Db	714 CGGCAGCGCTTCTCTGGAGAGGAGACACACGCCCTCGGACGGAGGTTCGCGAGCTA 770		
ALIGNMENT			
Pred. No.:	1.2e-11	Length:	855
Score:	19.00	Matches:	19
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	65.52%	Indels:	0
DB:	9	Gaps:	0
US-10-624-218-1 (1-29) x HSU06935 (1-855)			
Qy	1 ArgAlaAlaPheLeuGluLysGluAsnThrAlaLeuArgThrGluValAlaGluLeu 19		
Db	714 CGGCAGCGCTTCTCTGGAGAGGAGACACACGCCCTCGGACGGAGGTTCGCGAGCTA 770		
ALIGNMENT			
Pred. No.:	1.2e-11	Length:	855
Score:	19.00	Matches:	19
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	65.52%	Indels:	0
DB:	9	Gaps:	0
US-10-624-218-1 (1-29) x HSU06935 (1-855)			
Qy	1 ArgAlaAlaPheLeuGluLysGluAsnThrAlaLeuArgThrGluValAlaGluLeu 19		
Db	714 CGGCAGCGCTTCTCTGGAGAGGAGACACACGCCCTCGGACGGAGGTTCGCGAGCTA 770		
ALIGNMENT			
Pred. No.:	1.2e-11	Length:	855
Score:	19.00	Matches:	19
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	65.52%	Indels:	0
DB:	9	Gaps:	0
US-10-624-218-1 (1-29) x HSU06935 (1-855)			
Qy	1 ArgAlaAlaPheLeuGluLysGluAsnThrAlaLeuArgThrGluValAlaGluLeu 19		
Db	714 CGGCAGCGCTTCTCTGGAGAGGAGACACACGCCCTCGGACGGAGGTTCGCGAGCTA 770		
ALIGNMENT			
Pred. No.:	1.2e-11	Length:	855
Score:	19.00	Matches:	19
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	65.52%	Indels:	0
DB:	9	Gaps:	0
US-10-624-218-1 (1-29) x HSU06935 (1-855)			
Qy	1 ArgAlaAlaPheLeuGluLysGluAsnThrAlaLeuArgThrGluValAlaGluLeu 19		
Db	714 CGGCAGCGCTTCTCTGGAGAGGAGACACACGCCCTCGGACGGAGGTTCGCGAGCTA 770		
ALIGNMENT			
Pred. No.:	1.2e-11	Length:	855
Score:	19.00	Matches:	19
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	65.52%	Indels:	0
DB:	9	Gaps:	0
US-10-624-218-1 (1-29) x HSU06935 (1-855)			
Qy	1 ArgAlaAlaPheLeuGluLysGluAsnThrAlaLeuArgThrGluValAlaGluLeu 19		
Db	714 CGGCAGCGCTTCTCTGGAGAGGAGACACACGCCCTCGGACGGAGGTTCGCGAGCTA 770		
ALIGNMENT			
Pred. No.:	1.2e-11	Length:	855
Score:	19.00	Matches:	19
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	65.52%	Indels:	0
DB:	9	Gaps:	0
US-10-624-218-1 (1-29) x HSU06935 (1-855)			
Qy	1 ArgAlaAlaPheLeuGluLysGluAsnThrAlaLeuArgThrGluValAlaGluLeu 19		
Db	714 CGGCAGCGCTTCTCTGGAGAGGAGACACACGCCCTCGGACGGAGGTTCGCGAGCTA 770		
ALIGNMENT			
Pred. No.:	1.2e-11	Length:	855
Score:	19.00	Matches:	19
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	65.52%	Indels:	0
DB:	9	Gaps:	0
US-10-624-218-1 (1-29) x HSU06935 (1-855)			
Qy	1 ArgAlaAlaPheLeuGluLysGluAsnThrAlaLeuArgThrGluValAlaGluLeu 19		
Db	714 CGGCAGCGCTTCTCTGGAGAGGAGACACACGCCCTCGGACGGAGGTTCGCGAGCTA 770		
ALIGNMENT			
Pred. No.:	1.2e-11	Length:	855
Score:	19.00	Matches:	19
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	65.52%	Indels:	0
DB:	9	Gaps:	0
US-10-624-218-1 (1-29) x HSU06935 (1-855)			
Qy	1 ArgAlaAlaPheLeuGluLysGluAsnThrAlaLeuArgThrGluValAlaGluLeu 19		
Db	714 CGGCAGCGCTTCTCTGGAGAGGAGACACACGCCCTCGGACGGAGGTTCGCGAGCTA 770		
ALIGNMENT			
Pred. No.:	1.2e-11	Length:	855
Score:	19.00	Matches:	19
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	65.52%	Indels:	0
DB:	9	Gaps:	0
US-10-624-218-1 (1-29) x HSU06935 (1-855)			
Qy	1 ArgAlaAlaPheLeuGluLysGluAsnThrAlaLeuArgThrG		

DEFINITION	Human thyrotroph embryonic factor (TEF) mRNA, complete cds.	REFERENCE	1	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
ACCESSION	U44059	AUTHORS	Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.	
VERSION	U44059.1	TITLE	Kits, such as nucleic acid arrays, comprising a majority of humanexons or transcripts, for detecting expression and other uses thereof	
KEYWORDS		JOURNAL	Patent: WO 02068579-A 16197 06-SEP-2002; PE Corporation (NY) (US)	
SOURCE	Homo sapiens (human)	FEATURES	Location/Qualifiers	
ORGANISM	Homo sapiens	source	1..1003	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	/organism="Homo sapiens"		
REFERENCE	1 (bases 1 to 1000)	/mol_type="unassigned DNA"		
AUTHORS	Hunger, S.P., Li, S., Fall, M.Z., Naumovski, L. and Cleary, M.L.	/db_xref="taxon:9606"		
TITLE	The proto-oncogene HLF and the related basic leucine zipper protein TEF display highly similar DNA-binding and transcriptional regulatory properties			
JOURNAL	Blood 87 (11), 4607-4617 (1996)			
MEDLINE	96219638			
PUBMED	8639829			
REFERENCE	2 (bases 1 to 1000)	Alignment Scores:	1.38e-11	Length: 1003
AUTHORS	Hunger, S.P.	Pred. No.:	19.00	Matches: 19
JOURNAL	Submitted (29-DEC-1995) Stephen P. Hunger, Pediatrics, University of Colorado Health Sciences Center, UCHSC, Box C229, 4200 E 9th Ave, Denver, CO 80262, USA	Score:	100.00%	Conservative: 0
		Percent Similarity:	100.00%	Mismatches: 0
		Best Local Similarity:	65.52%	Indels: 0
		Query Match:	6	Gaps: 0
		DB:		
FEATURES	Location/Qualifiers	US-10-624-218-1 (1-29) x CQ730263 (1-1003)		
source	1..1000			
	/organism="Homo sapiens"			
	/mol_type="mRNA"			
	/db_xref="taxon:9606"			
	/chromosome="22"			
	/map="22q13.2"			
	/cell_line="ALL cell line HB11.19"			
gene	1..1000			
	/gene="TEF"			
CDS	37..948			
	/gene="TEF"			
	/note="TEF; similar to product encoded by Gallus gallus vbp gene, GenBank Accession number U09221; PAR-bZIP transcription factor"			
	/codon_start=1			
	/product="thyrotroph embryonic factor"			
	/protein_id="AA08497.1"			
	/db_xref="GI:1399386"			
	/translation="MSDAGGGKPPVPDQAGPGPGRAAGRLSGSPFLVLKKLME NPPRRARLDKEGKEKLEDEAAASTMAVASALMPTIWDKTIPIYDGESFHELYMDLD EFLLENGIPASPTLHAHLNLLPVALEKGEKASSTASPPSSSTAIFQPSVETVSTES SLEKRETFSPIDPNCVDFNFDPADLVLSVPGGELFNPGRKHPFAEDLKPQM IKKAKVFPDQKDEKVTWRKKNVAAKSRDARLKENQIITIRAAFLKENTALR TEVALRKEVGKCKTIVSKYETKYGL"			
ORIGIN				
Alignment Scores:	1.38e-11	Length:	1000	
Pred. No.:	19.00	Matches:	19	
Score:	100.00%	Conservative:	0	
Percent Similarity:	100.00%	Mismatches:	0	
Best Local Similarity:	65.52%	Indels:	0	
Query Match:	9	Gaps:	0	
DB:				
US-10-624-218-1 (1-29) x HSU44059 (1-1000)				
QY	1	ArgAlaAlaPheLeuGluIysGluAenThrAlaLeuArgThrGluValAlaGluLeu	19	
Db	826	CGGGCAGCCTTCCTCGGAGGAGAGACACAGCCCTCGGACGAGGTGCCGAGCTA	882	
RESULT 7				
CQ730263				
LOCUS	CQ730263	Sequence	16197 from Patent WO02068579.	
DEFINITION	CQ730263			
ACCESSION	CQ730263.1	GI:42303641		
VERSION				
KEYWORDS	Homo sapiens (human)			
SOURCE	Homo sapiens			
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			

Alignment Scores:	1.38e-11	Length:	1003	
Pred. No.:	19.00	Matches:	19	
Score:	100.00%	Conservative:	0	
Percent Similarity:	100.00%	Mismatches:	0	
Best Local Similarity:	65.52%	Indels:	0	
Query Match:	6	Gaps:	0	
DB:				
US-10-624-218-1 (1-29) x CQ730263 (1-1003)				
QY	1	ArgAlaAlaPheLeuGluIysGluAenThrAlaLeuArgThrGluValAlaGluLeu	19	
Db	829	CGGGCAGCCTTCCTCGGAGGAGAGACACAGCCCTCGGACGAGGTGCCGAGCTA	885	
RESULT 8				
GGU09221				
LOCUS	GGU09221	Gallus gallus White Leghorn beta/beta isoform of vitellogenin	1153 bp	mRNA linear VRT 30-NOV-1995
DEFINITION	GGU09221	binding protein (vbp) mRNA, complete cds.		
ACCESSION	U09221.1	GI:483937		
VERSION				
KEYWORDS	Gallus gallus (chicken)			
SOURCE	Gallus gallus			
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.			
REFERENCE	Burch, J.B. and Davis, D.L.			
AUTHORS	Alternative promoter usage and splicing options result in the differential expression of mRNAs encoding four isoforms of chicken VBP, a member of the PAR subfamily of bZIP transcription factors			
TITLE	Nucleic Acids Res. 22 (22), 4733-4741 (1994)			
JOURNAL	95075656			
MEDLINE	7984425			
PUBMED				
REFERENCE	2 (bases 1 to 1153)			
AUTHORS	Burch, J.B.			
TITLE	Direct Submission			
JOURNAL	Submitted (26-APR-1994) John B.E. Burch, Fox Chase Cancer Center, 7701 Burholme Ave., Philadelphia, PA 19111, USA			
FEATURES	Location/Qualifiers			
source	1..1153			
	/organism="Gallus gallus"			
	/mol_type="mRNA"			
	/strain="White Leghorn"			
	/db_xref="taxon:9031"			
	/cell_type="fibroblast"			
	/clone_lib="Vennstrom"			
	/dev_stages="embryonic day 10"			
	1..1153			
	/gene="vbp"			
	127..1008			
	/gene="vbp"			
	/codon_start=1			
	/product="vitellogenin binding protein (VBP), beta/beta isoform"			
	/protein_id="AAA82156.1"			
	/db_xref="GI:483938"			
gene				
CDS				

/translation="MSVCNAGGPAALDPPEVLKSLLEYSIPWTMTDKKKIKLE
DEAAASTMVASLMPPIWDKTIPIYDGESEHLEMDLDEFLLENGIPSSPHLDLQ
NPLMPVAKLEKPEASASTGSPVSSSTAVYQOEAASSTSPQNERNTSPDPDC
VEVEVFNPPADLVLSVPGGELFNPBKHFTEEDLKQPMIKKAKKVFVDPQKOE
KYWTRKNNVAKRSDRRARLKENQITIRAFLEKENTALRTEVAELRKEVGRCKNI
VSKYTRGPFDELSDSE"

ORIGIN

Alignment Scores:

Pred. No.: 1.56e-11 Length: 1153
Score: 19.00 Matches: 19
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 65.52% Indels: 0
DB: 5 Gaps: 0

US-10-624-218-1 (1-29) x GGU09221 (1-1153)

Qy 1 ArgAlaAlaPheLeuGluLysGluAsnThrAlaLeuArgThrGluValAlaGluLeu 19
Db 868 CGGCAGGCTTCCTTCGAGAAAGAGAAATACGGCCCTGAGACGGAGGTTCAGAGCTG 924

RESULT 9

AF194420 AF194420 1304 bp mRNA linear ROD 19-MAY-2000
LOCUS Mus musculus thymotroph embryonic factor (Tef) mRNA, complete cds.
ACCESSION AF194420
VERSION AF194420.1 GI:7939635

KEYWORDS

Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Krueger, D.A., Warner, E.A. and Dowd, D.R.
1 (bases 1 to 1304)

REFERENCE

Involvement of thymotroph embryonic factor in calcium-mediated

regulation of gene expression

J. Biol. Chem. 275 (19), 14524-14531 (2000)

20261592

10799536

2 (bases 1 to 1304)

Dowd, D.R., Krueger, D.A. and Warner, B.A.

Direct Submission

Submitted (12-OCT-1999) Pharmacology, Case Western Reserve

University School of Medicine, 10900 Euclid Ave., Cleveland, OH

44106-4965, USA

FEATURES

source

1. .1304 Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/strain="BALB/c"
/db_xref="taxon:10090"
/sex="male"
/tissue_type="brain"
/dev_stage="adult"
1. .1304
/gene="Tef"
91. .996
/gene="Tef"
/note="TEF; PAR-bZIP transcription factor"
/codon_start=1
/product="thymotroph embryonic factor"
/protein_id="AAF70830.1"
/db_xref="GI:7939636"

gene

CDS

/translation="MSDAGGKPPVPEPGAGRGSLGSPFLVLKLMENP
PRETRLDKEKGEKLEDESAASTWVASLMPPIWDKTIPIYDGESEHLEMDLDEFL
LENGIPSSPHLDLQNLPLVALEKESASSTASPPSSSTAIQFQSETSTESSL
EXERETPSIDPSCEVDVFNFPDADLVLSVPGGELFNPBKHFTEEDLKQPMIK
KAKKVFVDPQKOEKYWTRKNNVAKRSDRRARLKENQITIRAFLEKENTALRTE
VBLRKEVGRCKTIVSKYTKYGPL"

ORIGIN

Alignment Scores:

Pred. No.: 1.74e-11 Length: 1304
Score: 19.00 Matches: 19
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 65.52% Indels: 0
DB: 10 Gaps: 0

US-10-624-218-1 (1-29) x AF194420 (1-1304)

Qy 1 ArgAlaAlaPheLeuGluLysGluAsnThrAlaLeuArgThrGluValAlaGluLeu 19
Db 874 CGGCAGGCTTCCTTCGAGAAAGAGAAACACGCCCTCGGACGGAGGTTCGCGAGCTT 930

RESULT 10

AK091916

LOCUS

DEFINITION

AK091916

ACCESSION

AK091916.1 GI:21750394

VERSION

oligo capping; fis (full insert sequence).

KEYWORDS

Homo sapiens (human)

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1

Ota, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R.,

Wakamatsu, A., Hayashi, K., Sato, H., Nagai, K., Kimura, K., Makita, H.,

Sehara, M., Obayashi, M., Nishi, T., Shibahara, T., Tanaka, T.,

Ishii, S., Yamamoto, J., Saito, K., Kawai, Y., Isono, Y., Nakamura, Y.,

Nagahari, K., Murakami, K., Yasuda, T., Iwayanagi, T., Wagatsuma, M.,

Shiratori, A., Sudo, H., Hosoiri, T., Kaku, Y., Kodaira, H., Kondo, H.,

Sugawara, M., Takahashi, M., Kanda, K., Yokoi, T., Furuya, T.,

Kikkawa, E., Omura, Y., Abe, K., Kamihara, K., Katsuta, N., Sato, K.,

Tanikawa, M., Yamazaki, M., Ninomiya, K., Ishibashi, T., Yamashita, H.,

Murakawa, K., Fujimori, K., Tanai, H., Kimata, M., Watanabe, M.,

Hiraoka, S., Chiba, Y., Ishida, S., Ono, Y., Takiguchi, S., Watanabe, S.,

Yosida, M., Hotsuta, T., Kusano, J., Kanehori, K., Takahashi-Fujii, A.,

Hara, H., Tanase, T., Nomura, Y., Togiya, S., Komai, F., Hara, R.,

Takeuchi, K., Arita, M., Imose, N., Musashino, K., Yuuki, H., Oshima, A.,

Sasaki, N., Aotsuka, S., Yoshikawa, Y., Matsunawa, H., Ichihara, T.,

Shichata, N., Sano, S., Moriya, S., Momiyama, H., Satoh, N., Takami, S.,

Terashima, Y., Suzuki, O., Nakagawa, S., Senoh, A., Mizoguchi, H.,

Goto, Y., Shimizu, F., Wakebe, H., Hishigaki, H., Watanabe, T.,

Sugiyama, A., Takemoto, M., Kawakami, B., Yamazaki, M., Watanabe, K.,

Kunagai, A., Itakura, S., Fukuzumi, Y., Fujimori, Y., Komiyama, M.,

Tashiro, H., Tanigami, A., Fujiwara, T., Ono, T., Yamada, K., Fujii, Y.,

Ozaki, K., Hirao, M., Ohmori, Y., Kawabata, A., Hikiji, T., Kobatake, N.,

Inagaki, H., Ikema, Y., Okamoto, S., Okitani, R., Kawakami, T.,

Noguchi, S., Itoh, T., Shigeta, K., Senba, T., Matsumura, K.,

Nakajima, Y., Mizuno, T., Morinaga, M., Sasaki, M., Togashi, T.,

Oyama, M., Hata, H., Watanabe, M., Komatsu, T., Mizushima-Sugano, J.,

Satoh, T., Shirai, Y., Takahashi, Y., Nakagawa, K., Okumura, K.,

Nagase, T., Nomura, N., Kikuchi, H., Masuho, Y., Yamashita, R.,

Nakai, K., Yada, T., Nakamura, Y., Ohara, O., Isogai, T. and Sugano, S.

Complete sequencing and characterization of 21,243 full-length

human cDNAs

Nat. Genet. 36 (1), 40-45 (2004)

14702039

2

Tashiro, H., Yamazaki, M., Watanabe, K., Kumagai, A., Itakura, S.,

Fukuzumi, Y., Fujimori, Y., Komiyama, M., Sugiyama, T., Irie, R.,

Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y.,

Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H.,

Matsu, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K.,

Wagatsuma, M., Murakawa, K., Kanehori, K., Takahashi-Fujii, A.,

Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S.,

Nagahari, K., Masuho, Y., Nagai, K. and Isogai, T.

NEDO human cDNA sequencing project

Unpublished

3 (bases 1 to 2823)

Isogai, T. and Yamamoto, J.

Direct Submission

4

polya_signal
polya_site

ORIGIN

Alignment Scores:
Pred. No.: 4,48e-11 Length: 3808
Score: 19.00 Matches: 19
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 65.52% Indels: 0
DB: 9 Gaps: 0

US-10-624-218-1 (1-29) x HSM805945 (1-3808)

QY 1 ArgAlaAlaPheLeuGluLysGluAnthrAlaLeuArgThrGluValAlaGluLeu 19
|||||
DB 303 CGGGCAGCCTTCCTCGAGAAGAGAACACACAGCCCTCGCGACGGGTGCGCGAGCTA 359

RESULT 13
BC017689 4018 bp mRNA linear ROD 30-JUN-2004
LOCUS Mus musculus thymotroph embryonic factor, transcript variant 2,
DEFINITION mRNA (cDNA clone MGC:19233 IMAGE:4242534), complete cds.
ACCESSION BC017689
VERSION BC017689.1 GI:17389268
KEYWORDS MGC.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE
AUTHORS Strausberg,R.L., Peingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Schmen,C.N., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Schaeetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
Carninci,P., Frange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullaly,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahay,J., Helton,E., Kerteman,M., Madan,A., Rodriguez,S.,
Sanchez,A., Whitting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalls,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 4018)
Strausberg,R.
Direct Submission
Submitted (03-DEC-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgaps@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-shgc.stanford.edu>
Contact: (Dickson, Mark) mcd@paxil.stanford.edu

Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 25 Row: j Column: 24
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 23821037.
Location/Qualifiers
1. 4018
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="MGC:19233 IMAGE:4242534"
/tissue_type="kidney, normal. 5 month old male mouse."
/clone_lib="NCI CGAP_Kid14"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
1. 4018
/gene="Tef"
/db_xref="LocusID:21685"
/db_xref="MGI:98663"
4. 825
/gene="Tef"
/codon_start=1
/product="thymotroph embryonic factor, isoform 2"
/protein_id="AAH17689.1"
/db_xref="GI:17389269"
/db_xref="LocusID:21685"
/db_xref="MGI:98663"
/translation="MDMPEVLKSLLEHSLPWSKADKEKKEKLEDESAASAASTWAV
SASLMPPTDKTIPYDFSGFHELYMDLDFLENGIPASPTHLAQNLLLPVLEEGKE
SASLSTASPPSSSTAIFQSETSVSTSSLEKERETPSIDPCVEDVFNFPDPAFL
VLSSVPGGELFNPRKHFRAEDLKQPMIKKAKVFVPEQKDEKYWTRKKNNYAAK
RSRDARLKENQITIRAAFLKENTALTAEVLEKVEGCKCTIVSKYETKYGPL"

FEATURES
source
gene
CDS
ORIGIN

Alignment Scores:
Pred. No.: 4,69e-11 Length: 4018
Score: 19.00 Matches: 19
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 65.52% Indels: 0
DB: 10 Gaps: 0

US-10-624-218-1 (1-29) x BC017689 (1-4018)

QY 1 ArgAlaAlaPheLeuGluLysGluAnthrAlaLeuArgThrGluValAlaGluLeu 19
|||||
DB 703 CGGGCAGCCTTCCTCGAGAAGAGAACACACAGCCCTCGCGACGGGTGCGCGAGCTT 759

RESULT 14
AY540631 4064 bp mRNA linear ROD 12-APR-2004
LOCUS Mus musculus thymotroph embryonic factor alpha isoform (Tef) mRNA,
DEFINITION complete cds, alternatively spliced.
ACCESSION AY540631
VERSION AY540631.1 GI:42768793
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Eukaryota; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS Zhou,J., Hoggatt,A.M. and Herring,B.P.
TITLE Activation of the smooth muscle-specific telokin gene by thymotroph
embryonic factor (TEF)
J. Biol. Chem. 279 (16), 15929-15937 (2004)
PUBLISHED 14702338
REFERENCE
AUTHORS Zhou,J., Hoggatt,A.M. and Herring,B.P.
TITLE Direct Submission

JOURNAL	Submitted (30-JAN-2004) Cellular and Integrative Physiology, Indiana University School of Medicine, 635 Barnhill Drive, M2067, Indianapolis, IN 46202, USA
FEATURES	Location/Qualifiers
source	1. .4064 /organism="Mus musculus" /mol_type="mRNA" /db_xref="taxon:10090" /chromosome="15" /tissue lib="bladder"
gene	1. .4064 /gene="Tef"
CDS	143..1048 /gene="Tef" /note="transcription factor; alternatively spliced" /codon_start=1 /product="thyrotroph embryonic factor alpha isoform" /protein_id="AAS45599.1" /db_xref="GI:42768794" /translation="MSDAGGKKPPVPEQAGPGRAAGRLSGSPFLVKKLMENP PRETRDKKEKLEDESAASTMAVSASLMPPIWDKTIPIYDGESFHFLEYMDLDEF LLENGIPASPTHLAQNLLPVAELGKESASSSTASPSSSTAI FQSETVSTESSL EKERTSPIDPSCVEVDVNFPPADLVLSVPGGLFNPKRHRFAEDLKPQPMIK KAKKVFVPEQDEQDKYTRKKNVAAKSRDARLKENQITIRAAFLKENTALRTE VAELRKEVGKCKTIVSKYTKYGL"
ORIGIN	
Alignment Scores:	
Pred. No.:	4.74e-11 Length: 4064
Score:	19.00 Matches: 19
Percent Similarity:	100.00% Conservative: 0
Best Local Similarity:	100.00% Mismatches: 0
Query Match:	65.52% Indels: 0
DB:	10 Gaps: 0
US-10-624-218-1 (1-29) x AY540631 (1-4064)	
QY	1 ArgAlaAlaPheLeuGluysGluAsnThrAlaLeuArgThrGluValAlaGluLeu 19
Db	926 CGGCAGCGTTCCTGGAGAGGAGAACACAGCCCTCGCGAGGAGTTCGCGAGCTT 982
RESULT 15	
BC036982	
LOCUS	BC036982.1 GI:22477947
DEFINITION	Mus musculus thyrotroph embryonic factor, transcript variant 1, mRNA (cDNA clone MGC:46858 IMAGE:4976241), complete cds.
ACCESSION	BC036982
VERSION	BC036982
KEYWORDS	MGC.
SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS	1 (bases 1 to 4222) Klausner, R.D., Collins, F.S., Wagner, L., Shemen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, F.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullaby, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, M., Kettelman, M., Madan, A., Rodrigues, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Bickford, Y.S., Krzywinski, M.I., Skalska, U., Smalls, D.E., Schnerch, A., Schein, J.E., Jones, S.O. and Marra, M.A.
TITLE	Generation and initial analysis of more than 15,000 full-length
JOURNAL	human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A.	99 (26), 16899-16903 (2002)
12477932	
2 (bases 1 to 4222)	
Strausberg, R.	
Direct Submission	
Submitted (23-AUG-2002)	National Institutes of Health, Mammalian
Gene Collection (MGC),	Cancer Genomics Office, National Cancer
Institute, 31 Center Drive,	Room 11A03, Bethesda, MD 20892-2590,
USA	
NIH-MGC Project URL:	http://mgc.nci.nih.gov
Contact: MGC help desk	
Email: cgapbs@mail.nih.gov	
Tissue Procurement: Jeffrey Green M.D.	
cDNA Library Preparation: Life Technologies, Inc.	
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)	
DNA Sequencing by: National Institutes of Health Intramural	
Sequencing Center (NISC),	
Gaithersburg, Maryland;	
Web site: http://www.nisc.nih.gov/	
Contact: nisc.mgc@nih.gov	
Akhter, N., Ayale, K., Beckstrom-Sternberg, S.M., Benjamin, B.,	
Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,	
Dietrich, N.B., Granite, S., Guan, X., Gupta, J., Haghighi, P.,	
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,	
Maduro, Q.L., Masiello, C., Mastriani, B., Mastrian, S.D., McCloskey, J.C.,	
McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,	
Tsurgon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,	
Young, A., Zhang, L.-H. and Green, E.D.	
Clone distribution: MGC clone distribution information can be found	
through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov	
Series: IRAK Plate: 81 Row: g Column: 3	
This clone was selected for full length sequencing because it	
passed the following selection criteria: matched mRNA gi: 23821034.	
FEATURES	
source	1. .4222 /organism="Mus musculus" /mol_type="mRNA" /strain="FVB/N" /db_xref="taxon:10090" /clone="MGC:46858 IMAGE:4976241" /tissue_type="Mammary tumor. C3(1)-Tag model. Infiltrating ductal carcinoma. 5 month old virgin mouse." /clone_lib="NCI CGAP_Mam6" /lab_host="DH10B" /note="Vector: pCMV-SPORT6"
gene	1. .4222 /gene="Tef"
CDS	/db_xref="LocusID:21685" 121..1026 /gene="Tef" /codon_start=1 /product="thyrotroph embryonic factor, isoform 1" /protein_id="AAH36982.1" /db_xref="GI:22477948" /db_xref="LocusID:21685" /db_xref="MGI:98663" /translation="MSDAGGKKPPVPEQAGPGRAAGRLSGSPFLVKKLMENP PRETRDKKEKLEDESAASTMAVSASLMPPIWDKTIPIYDGESFHFLEYMDLDEF LLENGIPASPTHLAQNLLPVAELGKESASSSTASPSSSTAI FQSETVSTESSL EKERTSPIDPSCVEVDVNFPPADLVLSVPGGLFNPKRHRFAEDLKPQPMIK KAKKVFVPEQDEQDKYTRKKNVAAKSRDARLKENQITIRAAFLKENTALRTE VAELRKEVGKCKTIVSKYTKYGL"
ORIGIN	
Alignment Scores:	
Pred. No.:	4.9e-11 Length: 4222
Score:	19.00 Matches: 19
Percent Similarity:	100.00% Conservative: 0
Best Local Similarity:	100.00% Mismatches: 0
Query Match:	65.52% Indels: 0

DB: 10 Gaps: 0
US-10-624-218-1 (1-29) x BC036982 (1-4222)
Qy 1 ArgAlaAlaPheLeuGluLysGluAsnThrAlaLeuArgThrGluValAlaGluLeu 19
Db 904 CGGGCAGCCTTCTGGAGAGGAGAACACAGCCCTGCGGACGGAGGTTGCCGAGCTT 960

Search completed: July 27, 2005, 13:54:04
Job time : 1210.08 secs

This Page Blank (usptc)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 27, 2005, 10:05:45 ; Search time 232.403 Seconds
(without alignments)
738.685 Million cell updates/sec

Title: US-10-624-218-1

Perfect score: 29

Sequence: 1 RAAPFLEKENTALRTEVAEKEVGRCENI 29

Scoring table:

OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Word size: 1

Total number of hits satisfying chosen parameters: 8774627

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:
-MODEL=frame_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spoop_US10624218/runat_26072005_121435_3677/app_query.fasta_1.796
-DB=N Geneseq_16Dec04 -QFMT=fastap -SUFFIX=olip2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10624218 @CGN 1.1.1241 @runat_26072005_121435_3677 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THRRADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

N Geneseq_16Dec04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002as:*
7: Geneseqn2002bs:*
8: Geneseqn2003as:*
9: Geneseqn2003bs:*
10: Geneseqn2003cs:*
11: Geneseqn2003ds:*
12: Geneseqn2004as:*
13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	19	65.5	446	6	ABI99561 Mouse isc
2	19	65.5	999	13	Adg87162 Human tum
3	19	65.5	999	13	Adg87502 Human tum
4	19	65.5	1000	12	Adi82448 Human mod
5	19	65.5	1000	13	Acn40954 Tumour-as

6	19	65.5	4480	5	ABA16225	Abal6225 Human ner
7	19	65.5	5240	5	ABA16224	Abal6224 Human ner
c	8	31.0	217	12	ACH91032	Ach91032 Human gen
9	9	31.0	466	9	ACH35960	Ach35960 Human end
10	9	31.0	480	13	ADQ49232	Adq49232 Novel can
11	9	31.0	491	12	ACH88921	Ach88921 Human gen
12	9	31.0	546	12	ACH75221	Ach75221 Human gen
c	13	31.0	546	12	ACH77332	Ach77332 Human gen
14	9	31.0	1385	12	ADP13309	Adp13309 Renal cel
15	9	31.0	1385	13	ADR52783	Adr52783 Drug ther
16	9	31.0	1403	12	ADI82446	Adi82446 Human mod
17	9	31.0	1403	12	ADO19733	Ado19733 Human PRO
18	9	31.0	1403	13	ADP54858	Adp54858 Human PRO
19	9	31.0	1671	12	ADP72558	Adp72558 Renal tox
20	9	31.0	2678	12	ADQ87010	Adq87010 Human tum
21	9	31.0	2678	12	ADQ83706	Adq83706 Human tum
22	9	31.0	2678	12	ADQ85851	Adq85851 Human tum
c	23	31.0	5801	5	ABA18280	Abal18280 Human ner
24	8	27.6	57	12	ADL90017	Adl90017 Gluconoba
c	25	27.6	516	13	ACN58264	Acn58264 Cotton gy
c	26	27.6	546	13	ACN62069	Acn62069 Cotton gy
27	8	27.6	571	13	ACN60219	Acn60219 Cotton gy
28	8	27.6	1435	2	AAV16466	Aav16466 Nucleotid
29	8	27.6	1701	8	ACA47547	Ac47547 Prokaryot
30	8	27.6	1872	3	AAA47169	Aaa47169 DNA seque
31	8	27.6	1893	12	ADL90008	Adl90008 Gluconoba
32	8	27.6	2075	2	AAV16467	Aav16467 Nucleotid
33	8	27.6	2091	2	AAK18839	Aax18839 Streptoco
34	8	27.6	3207	10	ADE07583	Ado7583 Novel cod
35	8	27.6	3787	3	AAK59069	Aac59069 Human sec
36	8	27.6	5432	6	ABL67689	Abi67689 Oesophagu
37	8	27.6	5432	6	ABL64999	Abi64999 Lung canc
38	8	27.6	5432	6	ABK64387	Abk64387 Human ben
39	8	27.6	5773	12	ADI61744	Adi61744 Human cDN
40	8	27.6	6277	13	ADS10122	Ads10122 Human the
41	8	27.6	6282	4	AAK53013	Aak53013 Human pol
42	8	27.6	6286	4	AAK52029	Aak52029 Human pol
c	43	27.6	110000	12	ADQ97331_4	Continuation (5 of
44	7	24.1	197	3	AAC17847	Aac17847 Human sec
45	7	24.1	199	6	ABL79631	Abi79631 Human ova

ALIGNMENTS

RESULT 1
ABI99561
ID ABI99561 standard; CDNA; 446 BP.
XX
AC ABI99561;
XX
DT 07-MAR-2002 (first entry)
XX
DE Mouse ischaemic condition related cDNA sequence SEQ ID NO:564.
XX
KW Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;
KW vasospastic ischaemia; ischaemic condition; ischaemic disease; ss.
XX
OS Mus musculus.
XX
PN WO200188188-A2.
XX
PD 22-NOV-2001.
XX
PF 18-MAY-2001; 2001WO-JP004192.
XX
PR 18-MAY-2000; 2000JP-00145977.
XX
(UYNI-) UNIV NIHON SCHOOL JURIDICAL PERSON.
XX
PI Ishikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y;
XX
WPI; 2002-034733/04.
XX

PT Examining the ischemic condition (e.g. occlusive ischemia) by measuring
PT expression levels of particular genes defined in the specification or by
PT determining the expression profile of a gene group comprising these
XX genes.

XX Claim 2; Page 1498; 2690pp; English.

PS The present invention describes a method for examining ischaemic
CC conditions, comprising measuring the expression levels of particular
CC genes (I) in a test sample or determining the expression profile of a
CC gene group in the sample comprising genes selected from (I). The method
CC is useful for examining the ischaemic condition (e.g. compressive
CC ischaemia, occlusive ischaemia or vasospastic ischaemia) by measuring the
CC expression levels of particular genes (AB199202 to AB199912, encoding the
CC protein sequences in AB199202 to AB199912) or by determining the
CC expression profile of a gene group comprising these genes. The expression
CC levels or expression profiles produced by these genes are used as an
CC indicator when screening for ischaemic condition-improving drugs or
CC therapeutics for ischaemic diseases. AB199913 and AB199914 represent PCR
CC primers for a mouse ischaemic condition related sequence, which are used
XX in the exemplification of the present invention

SQ Sequence 446 BP; 109 A; 127 C; 129 G; 81 T; 0 U; 0 Other;

Alignment Scores: Length: 446
Pred. No.: 1.24e-10 Matches: 19
Score: 19.00
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 65.52% Indels: 0
DB: 6 Gaps: 0

US-10-624-218-1 (1-29) x AB199561 (1-446)

QY 1 ArgAlaAlaPheLeuGluLysGluAsnThrAlaLeuArgThrGluValAlaGluLeu 19
Db 88 CGGCAGCGTTCTTGGAGAGAGAGACACAGCCCTGCGGACGGAGTTGCGGAGCTT 144

RESULT 2
ADQ87162
ID ADQ87162 standard; cDNA; 999 BP.

XX AC ADQ87162;
XX 07-OCT-2004 (first entry)
XX Human tumour-associated antigenic target (TAT) cDNA sequence #4038.
XX human; tumour-associated antigenic target; TAT; cytostatic; gene therapy;
XX cancer; cell proliferative disorder; gene; ss.

XX Homo sapiens.
XX WO2004060270-A2.
XX 22-JUL-2004.
XX 15-OCT-2003; 2003WO-US029126.
XX 18-OCT-2002; 2002US-0418988P.
XX (GETH) GENENTECH INC.
XX (WUTD/) WU T D.
XX (ZHOU/) ZHOU Y.
XX Wu TD, Zhou Y;
XX WPI; 2004-534300/51.
XX New nucleic acid molecule and encoded polypeptide, for diagnosing,
XX preventing or treating cell proliferative disorders such as cancer.
XX Claim 1; SEQ ID NO 4038; 5504pp; English.

XX The present invention describes an isolated tumour-associated antigenic
CC target (TAT) nucleic acid comprising: (a) any of 4622 nucleotide
CC sequences (see SEQ ID NO:1 to 4622); (b) the full-length coding region of
CC (a); (c) the complement of (a) or (b); (d) a sequence that has 80%
CC sequence identity to (a)-(c); or (e) a sequence that hybridises to (a)-
CC (c). Also described: (1) an expression vector comprising the above
CC nucleic acid; (2) a host cell comprising the above expression vector; (3)
CC a process for producing a polypeptide; (4) an isolated polypeptide
CC comprising: (a) an amino acid sequence encoded by any of the above
CC nucleotide sequences; (b) an amino acid sequence encoded by the full-
CC length coding region of the above nucleotide sequences; or (c) a sequence
CC having at least 80% identical to (a) or (b); (5) a chimeric polypeptide
CC comprising the above polypeptide fused to a heterologous polypeptide; (6)
CC an isolated antibody that binds to the above polypeptide; (7) a process
CC for producing the antibody; (8) an isolated oligopeptide that binds to
CC the above polypeptide; (9) a tumour-associated antigenic target (TAT)
CC binding organic molecule that binds to the above polypeptide; (10) a
CC composition of matter comprising the above (chimeric) polypeptide,
CC antibody, oligopeptide or TAT binding organic molecule, in combination
CC with a carrier; (11) an article of manufacture comprising a container and
CC the composition of matter contained within the container; (12) methods of
CC inhibiting the growth of a cell that expresses the above protein, where
CC the growth of the cell is at least in part dependent upon a growth
CC potentiating effect of the above protein; (13) a method of
CC therapeutically treating a mammal having a cancerous tumour comprising
CC cells that express the above protein; (14) a method of determining the
CC presence of a protein in a sample suspected of containing the protein in a
CC described above; (15) methods of diagnosing the presence of a tumour in a
CC mammal; (16) a method for treating or preventing a cell proliferative or
CC disorder associated with increased expression or activity of the above
CC protein; and (17) a method of binding an antibody, oligopeptide or
CC organic molecule to a cell that expresses the protein described above.
CC The TAT sequences have cytostatic activities, and can be used in gene
CC therapy. The composition and methods are useful for diagnosing,
CC preventing or treating cancer. The composition is also used for preparing
CC a medicament for the therapeutic treatment or diagnostic detection of a
CC cell proliferative disorder or cancer. The present sequence represents a
CC human TAT cDNA sequence from the present invention.

XX Sequence 999 BP; 231 A; 311 C; 307 G; 150 T; 0 U; 0 Other;

Alignment Scores: Length: 999
Pred. No.: 2.56e-10 Matches: 199
Score: 19.00
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 65.52% Indels: 0
DB: 13 Gaps: 0

US-10-624-218-1 (1-29) x ADQ87162 (1-999)

QY 1 ArgAlaAlaPheLeuGluLysGluAsnThrAlaLeuArgThrGluValAlaGluLeu 19
Db 825 CGGCAGCGTTCTTGGAGAGAGAGACACAGCCCTGCGGACGGAGTTGCGGAGCTA 881

RESULT 3
ADQ87502
ID ADQ87502 standard; cDNA; 999 BP.

XX AC ADQ87502;
XX 07-OCT-2004 (first entry)
XX Human tumour-associated antigenic target (TAT) cDNA sequence #4380.
XX human; tumour-associated antigenic target; TAT; cytostatic; gene therapy;
XX cancer; cell proliferative disorder; gene; ss.

XX Homo sapiens.
XX WO2004060270-A2.
XX

PD 22-JUL-2004.
XX
XX
PF 15-OCT-2003; 2003WO-US029126.
XX
XX 18-OCT-2002; 2002US-041898BP.
XX
XX (GETH) GENENTECH INC.
PA (WUTD/) WU T D.
PA (ZHOU/) ZHOU Y.
XX
XX Wu TD, Zhou Y;
PI WPI; 2004-534300/51.
XX
XX New nucleic acid molecule and encoded polypeptide, for diagnosing,
PT preventing or treating cell proliferative disorders such as cancer.
XX
XX Claim 1; SEQ ID NO 4380; 5504pp; English.
XX
XX The present invention describes an isolated tumour-associated antigenic
CC target (TAT) nucleic acid comprising: (a) any of 4622 nucleotide
CC sequences (see SEQ ID NO:1 to 4622); (b) the full-length coding region of
CC (a); (c) the complement of (a) or (b); (d) a sequence that has 80%
CC sequence identity to (a)-(c); or (e) a sequence that hybridises to (a)-
CC (c). Also described: (1) an expression vector comprising the above
CC nucleic acid; (2) a host cell comprising the above expression vector; (3)
CC a process for producing a polypeptide; (4) an isolated polypeptide
CC comprising: (a) an amino acid sequence encoded by any of the above
CC nucleotide sequences; (b) an amino acid sequence encoded by the full-
CC length coding region of the above nucleotide sequences; or (c) a sequence
CC having at least 80% identical to (a) or (b); (5) a chimeric polypeptide
CC comprising the above polypeptide fused to a heterologous polypeptide; (6)
CC an isolated antibody that binds to the above polypeptide; (7) a process
CC for producing the antibody; (8) an isolated oligopeptide that binds to
CC the above polypeptide; (9) a tumour-associated antigenic target (TAT)
CC binding organic molecule that binds to the above polypeptide; (10) a
CC composition of matter comprising the above (chimeric) polypeptide,
CC antibody, oligopeptide or TAT binding organic molecule, in combination
CC with a carrier; (11) an article of manufacture comprising a container and
CC the composition of matter contained within the container; (12) methods of
CC inhibiting the growth of a cell that expresses the above protein, where
CC the growth of the cell is at least in part dependent upon a growth
CC potentiating effect of the above protein; (13) a method of
CC therapeutically treating a mammal having a cancerous tumour comprising
CC cells that express the above protein; (14) a method of determining the
CC presence of a protein in a sample suspected of containing the protein
CC described above; (15) methods of diagnosing the presence of a tumour in a
CC mammal; (16) a method for treating or preventing a cell proliferative
CC disorder associated with increased expression or activity of the above
CC protein; and (17) a method of binding an antibody, oligopeptide or
CC organic molecule to a cell that expresses the protein described above.
CC The TAT sequences have cytostatic activities, and can be used in gene
CC therapy. The composition and methods are useful for diagnosing,
CC preventing or treating cancer. The composition is also used for preparing
CC a medicament for the therapeutic treatment or diagnostic detection of a
CC cell proliferative disorder or cancer. The present sequence represents a
CC human TAT cDNA sequence from the present invention.
XX
SQ Sequence 999 BP; 231 A; 311 C; 307 G; 150 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2,56e-10 Length: 199
Score: 19.00 Matches: 19
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 65.52% Indels: 0
DB: 13 Gaps: 0

US-10-624-218-1 (1-29) x ADQ87502 (1-999)

QY 1 ArgAlaAlaPheLeuGluLysGluAsnThrAlaLeuArgThrGluValAlaGluLeu 19
DB 825 CGGGCAGCCTTCCTCGAGAAGGAGACACAGCCCTCGGACGAGGTGGCCGAGCTA 881

RESULT 4
ADI82448
ID ADI82448 standard; DNA; 1000 BP.
XX
XX AC ADI82448;
XX
XX 22-APR-2004 (first entry)
XX
XX Human modifier of p21 (MP21) gene sequence SeqID14.
XX
XX p21 pathway modulating agent; assay system; MP21; cytostatic;
KW MP21 protein activity modulator; cancer; genetically modified animal;
KW human; gene; ds.
XX
XX Homo sapiens.
XX
XX WO2004005486-A2.
XX
XX 15-JAN-2004.
XX
XX 09-JUL-2003; 2003WO-US021510.
XX
XX 10-JUL-2002; 2002US-0394795P.
PR 07-AUG-2002; 2002US-0401739P.
PR 16-SEP-2002; 2002US-0411010P.
PR 30-DEC-2002; 2002US-0437158P.
XX
XX (EXEL-) EXELIXIS INC.
XX
XX Francis-Lang H, Friedman L, Kidd T, Roche S, Joo DM, Lickteig K;
PI Amundsen CD, Hai B, Zhang H, Adamkewicz JI, Hammonds GR;
XX
XX WPI: 2004-091358/09.
DR P-PSDB; ADI82509.
XX
XX Identifying a candidate p21 pathway modulating agent, useful for treating
PT a disease such as cancer, comprises contacting an assay system comprising
PT a MP21 polypeptide or nucleic acid with a test agent.
XX
XX Example 2; SEQ ID NO 14; 392pp; English.
XX
XX This invention relates to a novel candidate p21 pathway modulating agent
CC by contacting an assay system comprising an MP21 (modifier of p21)
CC polypeptide or nucleic acid with a test agent, where in the absence of
CC the test agent the system provides a reference activity and detecting a
CC test agent-biased activity of the assay system. The invention may be
CC useful for the production of compounds with a cytostatic activity through
CC modulation of MP21 protein activity. The MP21 polypeptide or nucleic acid
CC can be used for identifying MP21 modulating agents useful as therapeutic
CC targets for diagnosing cancer or treating disorders associated with
CC defective or impaired p21 and/or MP21 function. MP21 modulating agents
CC are useful in diagnosis, therapy, for example treating cancer, and
CC pharmaceutical development. The genetically modified animals may be used
CC for in vivo assays to test for activity of a candidate p21 modulating
CC agent, or to further assess the role of MP21 in a p21 pathway process.
CC The present sequence is that of a human MP21 gene which is an orthologue
CC of a Drosophila p21 modifier and which was used in the exemplification of
CC the invention.
XX
SQ Sequence 1000 BP; 231 A; 312 C; 307 G; 150 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2,56e-10 Length: 1000
Score: 19.00 Matches: 19
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 65.52% Indels: 0
DB: 12 Gaps: 0

US-10-624-218-1 (1-29) x ADI82448 (1-1000)

QY 1 ArgAlaAlaPheLeuGluLysGluAsnThrAlaLeuArgThrGluValAlaGluLeu 19

Db 826 CGGCGAGCTTCTCTGGAGAGGAGAACACACAGCCCTGCGGACGGAGGTGGCCGAGCTA 882

RESULT 5
ACN40954
ID ACN40954 standard; cDNA; 1000 BP.
XX
XX
AC ACN40954;
XX
DT 18-NOV-2004 (first entry)
XX
DE Tumour-associated antigenic target (TAT) cDNA DNA326982, SEQ ID NO:6077.
XX
KW Tumour-associated antigenic target; TAT; human; overexpression; cancer;
KW tumour; diagnosis; cell proliferative disorder; breast cancer;
KW colorectal cancer; lung cancer; ovarian cancer; liver cancer;
KW central nervous system cancer; bladder cancer; pancreatic cancer;
KW cervical cancer; melanoma; leukaemia; hybridisation probe;
KW chromosome identification; chromsome mapping; gene mapping;
KW gene therapy; cytostatic; gene; ss.
XX
OS Homo sapiens.
XX
XX WO2004030615-A2.
XX
PD 15-APR-2004.
XX
XX 29-SEP-2003; 2003WO-US028547.
XX
PF
XX
XX 02-OCT-2002; 2002US-0414971P.
XX
XX (GETH) GENENTECH INC.
XX
XX Wu TD, Zhang Z, Zhou Y;
XX
XX WPI: 2004-347921/32.
XX P-PSDB; ABM82367.
XX
XX New tumor-associated antigenic target polypeptides and nucleic acids,
PT useful in preparing a medicament for treating or detecting a
PT proliferative disorder, e.g. breast, lung, colorectal, ovarian or
PT prostate cancer or tumor.
XX
XX Claim 1; SEQ ID NO 6077; 7273pp; English.
XX
XX The invention relates to human tumour-associated antigenic target (TAT)
CC polypeptides, and their related nucleic acids. The TAT polypeptides are
CC overexpressed in cancer tissues compared to normal tissues, and may thus
CC serve as effective targets for the diagnosis and treatment of cancer in
CC mammals. The invention also relates to nucleic acid and polypeptide
CC sequences at least 80% identical to the TAT nucleic acids and
CC polypeptides; expression vectors and host cells comprising a TAT nucleic
CC acid; an antibody specific for a TAT polypeptide; a peptide or organic
CC molecule which binds to a TAT polypeptide; fusion proteins comprising a
CC TAT polypeptide; and methods and compositions for the treatment of
CC diagnosis of cancer in mammals. TAT polypeptides, nucleic acids,
CC antibodies, antagonists, binding molecules and compositions are useful
CC for diagnosing or treating a cell proliferative disorder associated with
CC increased TAT expression, particularly cancers such as breast cancer,
CC colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder
CC cancer, pancreatic cancer, cervical cancer, cancers of the central
CC nervous system, melanoma and leukaemia. TAT nucleic acids may further be
CC used as hybridisation probes, in chromosome and gene mapping, in
CC chromosome identification and in gene therapy. The present sequence
CC represents a TAT nucleic acid of the invention
XX
SQ Sequence 1000 BP; 231 A; 312 C; 307 G; 150 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2 56e-10 Length: 1000
Score: 19.00 Matches: 19
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0

Query Match: 65.52% Indels: 0
DB: 13 Gaps: 0

US-10-624-218-1 (1-29) x ACN40954 (1-1000)

QY 1 ArgAlaAlaPheLeuGluLysGluAsnThrAlaLeuArgThrGluValAlaGluLeu 19
DB 826 CGGCGAGCTTCTCTGGAGAGGAGAACACAGCCCTGCGGACGGAGGTGGCCGAGCTA 882

RESULT 6
ABAI6225
ID ABAI6225 standard; DNA; 4480 BP.
XX
XX ABAI6225;
XX
DT 23-JAN-2002 (first entry)
XX
DE Human nervous system related polynucleotide SEQ ID NO 8556.
XX
XX Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;
KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;
KW antiparkinsonian; antisickling; antianaemic; antiarthritic; cancer;
KW antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
KW antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;
KW antiparasitic; cardiant; immune disorder; cardiovascular disorder;
KW neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.
XX
XX Homo sapiens.
XX
XX WO200159063-A2.
XX
XX 16-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US001334.
XX
XX 31-JAN-2000; 2000US-0179065P.
XX 04-FEB-2000; 2000US-0180828P.
XX 24-FEB-2000; 2000US-0184664P.
XX 02-MAR-2000; 2000US-0186350P.
XX 16-MAR-2000; 2000US-0189874P.
XX 17-MAR-2000; 2000US-0190076P.
XX 18-APR-2000; 2000US-0198123P.
XX 19-MAY-2000; 2000US-0205515P.
XX 07-JUN-2000; 2000US-0209467P.
XX 28-JUN-2000; 2000US-0214886P.
XX 30-JUN-2000; 2000US-0215135P.
XX 07-JUL-2000; 2000US-0216647P.
XX 07-JUL-2000; 2000US-0216880P.
XX 11-JUL-2000; 2000US-0217487P.
XX 11-JUL-2000; 2000US-0217496P.
XX 14-JUL-2000; 2000US-0218290P.
XX 26-JUL-2000; 2000US-0220963P.
XX 26-JUL-2000; 2000US-0220964P.
XX 14-AUG-2000; 2000US-0224518P.
XX 14-AUG-2000; 2000US-0224519P.
XX 14-AUG-2000; 2000US-0225213P.
XX 14-AUG-2000; 2000US-0225214P.
XX 14-AUG-2000; 2000US-0225266P.
XX 14-AUG-2000; 2000US-0225267P.
XX 14-AUG-2000; 2000US-0225268P.
XX 14-AUG-2000; 2000US-0225270P.
XX 14-AUG-2000; 2000US-0225447P.
XX 14-AUG-2000; 2000US-0225757P.
XX 14-AUG-2000; 2000US-0225758P.
XX 14-AUG-2000; 2000US-0225759P.
XX 18-AUG-2000; 2000US-0226279P.
XX 22-AUG-2000; 2000US-0226681P.
XX 22-AUG-2000; 2000US-0226868P.
XX 22-AUG-2000; 2000US-0227182P.
XX 23-AUG-2000; 2000US-0227009P.
XX 30-AUG-2000; 2000US-0228924P.
XX 01-SEP-2000; 2000US-0229287P.
XX 01-SEP-2000; 2000US-0229343P.


```
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229511P.
PR 06-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0231415P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0233977P.
PR 14-SEP-2000; 2000US-0233988P.
PR 14-SEP-2000; 2000US-0233999P.
PR 14-SEP-2000; 2000US-0234000P.
PR 14-SEP-2000; 2000US-0234011P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 26-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 20-OCT-2000; 2000US-0242221P.
PR 01-NOV-2000; 2000US-0244617P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246612P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.

PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250391P.
PR 01-DEC-2000; 2000US-0251160P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.

XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-541565/60.
XX
XX Nucleic acids encoding 3224 human nervous system antigen polypeptides,
XX useful for preventing, diagnosing and/or treating nervous system cancers
XX and metastases.
XX
XX Disclosure; SEQ ID NO 8556; 1701pp + Sequence Listing; English.
XX
XX The invention relates to novel genes (ABAI1004-ABA21534) and proteins
XX (ABBI14678-ABBI18001) useful for preventing, treating or ameliorating
XX medical conditions e.g. by protein or gene therapy. The genes are
XX isolated from a range of human tissues disclosed in the specification.
XX The nucleic acids, proteins, antibodies and (ant)agonists are useful in
XX the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
XX ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
XX breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
XX disorders e.g. Addison's disease, allergies, diabetes mellitus, Crohn's disease,
XX anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
XX multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
XX cardiovascular disorders such as myocardial ischaemia; (d) wound healing
XX ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
XX infectious diseases such as viral, bacterial, fungal and parasitic
XX infections. Note: The sequence data for this patent did not form part of
XX the printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 4480 BP; 913 A; 1195 C; 1214 G; 1158 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 9,96E-10 Length: 4480
Score: 19.00 Matches: 19
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 65.52% Indels: 0
DB: 5 Gaps: 0

US-10-624-218-1 (1-29) x ABA16225 (1-4480)

QY 1 ArgAlaAlaPheLeuGluLysGluAsnThrAlaLeuArgThrGluValAlaGluLeu 19
Db 1745 CGGGCAGCCTTCTCTGGAGAGGAGACACAGCCTCGCGAGGTGGCCGAGCTA 1801

RESULT 7
```

Thu Jul 28 11:34:40 2005

ABAI6224
ID ABA16224 standard; DNA; 5240 BP.
XX
AC ABA16224;
XX
DT 23-JAN-2002 (first entry)
XX
XX Human nervous system related polynucleotide SEQ ID NO 8555.
XX
XX Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;
KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;
KW antiparkinsonian; antischlicking; antianaemic; antiarthritic; cancer;
KW antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
KW antiallergic; antidiabetic; antiulcer; anticonvulsant; antifungal;
KW antiparasitic; cardiant; immune disorder; cardiovascular disorder;
KW neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.
XX
XX Homo sapiens.
XX
XX WO200159063-A2.
XX
XX 16-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US001334.
XX
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-024617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.

PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250391P.
PR 01-DEC-2000; 2000US-0251160P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251989P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Barash SC, Ruben SM;
PI WPI; 2001-541565/60.
DR Nucleic acids encoding 3224 human nervous system antigen polypeptides,
PT useful for preventing, diagnosing and/or treating nervous system cancers
PT and metastases.
XX Disclosure; SEQ ID NO: 8555; 1701pp + Sequence Listing; English.
PS The invention relates to novel genes (ABAI1004-ABAI1534) and proteins
XX (ABAI4678-ABAI8001) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in
CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
CC ovarian cancer and other cancers of the adrenal gland, bone, marrow,
CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
CC cardiovascular disorders such as myocardial ischaemia; (d) wound healing
CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
CC infectious diseases such as viral, bacterial, fungal and parasitic
CC infections. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 5240 BP; 1079 A; 1354 C; 1452 G; 1355 T; 0 U; 0 Other;

Alignment Scores:
Pred. NO.: 1.15e-09 Length: 5240
Score: 19.00 Matches: 19
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 65.52% Indels: 0
DB: 5 Gaps: 0

US-10-624-218-1 (1-29) x ABAI16224 (1-5240)
QY 1 ArgAlaAlaPheLeuGluLysGluAenThrAlaLeuArgThrGluValAlaGluLeu 19
DB 1745 CGGCGAGCCTTCCTGGAGAAGGAGACACAGCCCTCGCAGCGAGGTGGCCGAGCTA 1801

RESULT 8
ACH91032/c
ID ACH91032 standard; DNA; 217 BP.
XX
XX ACH91032;
AC
XX
DT 29-JUL-2004 (first entry)
DE Human genome derived single exon probe #24227.
XX
XX Human; probe; ss; gene expression; single exon probe; microarray;

KW alternative splicing event; genomic alteration.
XX Homo sapiens.
XX US2003194704-A1.
XX 16-OCT-2003.
XX
XX 03-APR-2002; 2002US-00029386.
XX 03-APR-2002; 2002US-00029386.
XX (PENN/) PENN S G.
XX (RANK/) RANK D R.
XX (HANZ/) HANZEL D K.
XX Penn SG, Rank DR, Hanzel DK;
XX WPI; 2004-119264/12.
XX
XX New human genome-derived single exon nucleic acid probes useful for human
PT gene expression analysis, for identifying or characterizing alternative
PT splicing events, for assessing genomic alterations or as tools for
PT surveying tissues.
XX
PS Claim 1; SEQ ID NO 24227; 80pp; English.
XX
CC The invention relates to a nucleic acid probe for measuring human gene
CC expression, comprising any of the 27,400 fully defined nucleotide
CC sequences in the specification, or their complements or fragments, and
CC encoding at least 8 amino acids of any of the 6888 amino acid sequences
CC fully defined in the specification. The probe is a single exon probe that
CC hybridises under high stringency conditions to a nucleic acid molecule
CC expressed in human cells or tissues. Also included are a spatially-
CC addressable set of single exon nucleic acid probes for measuring human
CC gene expression (comprising a plurality of single exon nucleic acid
CC probes cited above, where each of the plurality of probes is separately
CC and addressably isolatable or amplifiable from the plurality), a single
CC exon microarray for measuring human gene expression, a method of
CC measuring human gene expression, a vector comprising the single exon
CC probe cited above, an ORF-encoded peptide comprising at least 8
CC contiguous amino acids of any of the above-mentioned amino acid
CC sequences (optionally with conservative amino acid substitutions), an
CC isolated antibody that binds specifically to a peptide cited above,
CC a method of selling and/or licensing single exon probes or microarrays to
CC a customer desiring to measure gene expression, a method of providing
CC human gene expression data by subexpression, and a computer-readable
CC storage medium which contains a database having a plurality of records
CC (each record including data on the expression of a single exon probe
CC cited above. The probe, methods and apparatus are useful in gene
CC expression analysis. The probes may be used as tools for surveying
CC tissues to detect the presence of expressed messages that contain their
CC specific exon, or in constructing genome-derived single exon microarrays.
CC In addition, the probes are used in identifying and characterising
CC alternative splicing events, in detecting and characterising gross
CC alterations in the genomic locus that includes their exon, in assessing
CC smaller genomic alterations, in priming the synthesis of nucleic acids,
CC or in expressing the ORF-encoded peptide. The present sequence is a human
CC single exon probe of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20030194704
XX
SQ Sequence 217 BP; 28 A; 77 C; 66 G; 46 T; 0 U; 0 Other;

Alignment Scores:
Pred. NO.: 1.21 Length: 217
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 31.03% Indels: 0
DB: 12 Gaps: 0

US-10-624-218-1 (1-29) x ACH91032 (1-217)

QY 1 ArgAlaAlaPheLeuGluLysGluAsn 9
DB 123 CGGCGCGCCTTCCTGGAGAGGAGAAC 97

RESULT 9

ACH35960
ID ACH35960 standard; cDNA; 466 BP.

XX AC ACH35960;

XX DT 13-OCT-2003 (first entry)

XX DE Human endothelial cell cDNA #4093.

XX KW Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
genome mapping; biodiversity; genetic disorder.

XX OS Homo sapiens.

XX PN US2003073623-A1.

XX PD 17-APR-2003.

XX PF 30-JUL-2001; 2001US-00918995.

XX PR 30-JUL-2001; 2001US-00918995.

XX PA (DRMA/) DRMANAC R T.

XX PA (LABA/) LABAT I.

XX PA (STAC/) STACHE-CRAIN B.

XX PA (DICK/) DICKSON M C.

XX PA (JONE/) JONES L W.

XX PI Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;

XX DR WPI; 2003-615964/58.

XX PT New polynucleotide sequences obtained from various cDNA libraries, useful
as hybridization probes, as oligomers for PCR, for chromosome and gene
mapping, in the recombinant production of protein, or in generating
antisense DNA or RNA.

PS Claim 1; SEQ ID NO 23172; 44pp; English.

XX The invention relates to an isolated polynucleotide comprising any one of
38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
determined by the technique of SBH (sequencing by hybridisation). Also
included is a purified polypeptide comprising a sequence corresponding to
a reading frame of the novel polynucleotide. The nucleic acid sequences
are useful in diagnostics as expressed sequence tags (EST) for
identifying expressed genes or for physical mapping of the human genome,
in forensics, in assessing biodiversity, or in identifying mutations
responsible for genetic disorders and other traits. The nucleotide
sequences are also useful as hybridisation probes, as oligomers for PCR,
for chromosome and gene mapping, in the recombinant production of
protein, or in generating antisense DNA or RNA. The purified polypeptide
is useful for generating antibodies specific for it. The present sequence
is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
for this patent did not form part of the printed specification, but was
obtained in electronic format directly from USPTO at
seqdata.uspto.gov/sequence.html?DocID=20030073623

XX SQ Sequence 466 BP; 114 A; 133 C; 129 G; 90 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2-42	Length:	466
Score:	9.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	31.03%	Indels:	0
DB:	9	Gaps:	0

US-10-624-218-1 (1-29) x ACH35960 (1-466)

QY 1 ArgAlaAlaPheLeuGluLysGluAsn 9
DB 407 CGGCGCGCCTTCCTGGAGAGGAGAAC 433

RESULT 10

ADQ49232

XX ID ADQ49232 standard; DNA; 480 BP.

XX AC ADQ49232;

XX DT 21-OCT-2004 (first entry)

XX DE Novel canine microarray-related DNA sequence SeqID534.

XX KW canine microarray; drug screening; toxicity assay;
environmental pollutant; cellular response; gene expression profile;
toxic response; liver necrosis; fatty liver disease;
protein adduct formation; hepatitis; dog; ds.

XX OS Canis familiaris.

XX PN WO2004063324-A2.

XX PD 29-JUL-2004.

XX PF 05-MAY-2003; 2003WO-US013853.

XX PR 03-MAY-2002; 2002US-0377240P.

XX PA (GENE-) GENE LOGIC INC.

XX PA (PFIZ) PFIZER PROD INC.

XX PI Diggins JC, Porter M, Wei T;

XX DR WPI; 2004-561890/54.

XX PT New isolated nucleic acid molecule, useful for drug screening and
toxicity assays or for assessing the impact, including toxicity, of a
compound, pharmaceutical agent or environmental pollutant on a cell or
living organism.

PS Claim 1; SEQ ID NO 534; 41pp; English.

XX This invention is related to a novel isolated canine nucleic acid
sequences and the construction of canine microarrays containing a
significant portion of the canine genome. The isolated canine nucleic
acid sequences of the invention may be useful for drug screening and
toxicity assays. The invention is therefore useful for assessing the
impact, including toxicity, of a compound, pharmaceutical agent or
environmental pollutant on a cell or living organism. The methods are
useful for detecting genes that are up- or down-regulated in canines in a
disease state. The sequences are useful as diagnostic agents or markers
to detect a cellular response in a sample individually or as part of a
gene expression profile. It is also useful as a target for agents that
modulate gene expression or activity. The database is useful for the cell
producing electronic Northern that allow the user to determine the cell
type or tissue in which a given gene is expressed and to allow
determination of the abundance or expression level of a given gene in a
particular tissue or cell. The methods are useful for determining the
similarity of a toxic response to one or more individual compounds. The
methods are useful for predicting at least one toxic response or the
likelihood that a compound or test agent will induce various specific
pathologies such as those of the liver (liver necrosis, fatty liver
disease, protein adduct formation or hepatitis), those of the kidney,
heart, brain or testes, or other pathologies associated with at least one
of the toxins. The methods are also useful for predicting or elucidating
the potential cellular pathways influenced, induced or modulated by the
compound or test agent due to the similarity of the expression profile
compared to the profile induced by a known toxin. The present sequence is
that of a canine DNA sequence which was claimed for use during the

```
CC production of a canine microarray of the invention.
SQ Sequence 480 BP; 100 A; 151 C; 146 G; 77 T; 0 U; 6 Other;

Alignment Scores:
Pred. No.: 2.49 Length: 480
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 31.03% Indels: 0
DB: 13 Gaps: 0

US-10-624-218-1 (1-29) x ADQ49232 (1-480)
QY 1 ArgAlaAlaPheLeuGlutylsGluAen 9
DB 142 CGGGCGGCTTCTCTGGAGAGGAGAAC 168

RESULT 11
ACH88921
ID ACH88921 standard; DNA; 491 BP.
XX
AC ACH88921;
XX
DT 29-JUL-2004 (first entry)
XX
DE Human genome derived single exon probe #22116.
XX
KW Human; probe; ss; gene expression; single exon probe; microarray;
KW alternative splicing event; genomic alteration.
XX
OS Homo sapiens.
XX
PN US2003194704-A1.
XX
PD 16-OCT-2003.
XX
PF 03-APR-2002; 2002US-00029386.
XX
PR 03-APR-2002; 2002US-00029386.
XX
PA (PENN/) PENN S G.
PA (RANK/) RANK D R.
PA (HANZ/) HANZEL D K.
XX
PI Penn SG, Rank DR, Hanzel DK;
XX
XX WPI; 2004-119264/12.
XX
PT New human genome-derived single exon nucleic acid probes useful for human
PT gene expression analysis, for identifying or characterizing alternative
PT splicing events, for assessing genomic alterations or as tools for
PT surveying tissues.
XX
PS Claim 1; SEQ ID NO 22116; 80pp; English.
XX
XX
CC The invention relates to a nucleic acid probe for measuring human gene
CC expression, comprising any of the 27,400 fully defined nucleotide
CC sequences in the specification, or their complements or fragments, and
CC encoding at least 8 amino acids of any of the 688 amino acid sequences
CC fully defined in the specification. The probe is a single exon probe that
CC hybridises under high stringency conditions to a nucleic acid molecule
CC expressed in human cells or tissues. Also included are a spatially-
CC addressable set of single exon nucleic acid probes for measuring human
CC gene expression (comprising a plurality of single exon nucleic acid
CC probes cited above, where each of the plurality of probes is separately
CC and addressably isolatable or amplifiable from the plurality), a single
CC exon microarray for measuring human gene expression, a method of
CC measuring human gene expression, a vector comprising the single exon
CC probe cited above, an ORF-encoded peptide comprising at least 8
CC contiguous amino acids of any of the above-mentioned amino acid
CC sequences (optionally with conservative amino acid substitutions), an
CC isolated antibody that binds specifically to a peptide cited above,
```

```
CC methods of selling and/or licensing single exon probes or microarrays to
CC a customer desiring to measure gene expression, a method of providing
CC human gene expression data by subscription, and a computer-readable
CC storage medium which contains a database having a plurality of records
CC (each record including data on the expression of a single exon probe
CC cited above. The probe, methods and apparatus are useful in gene
CC expression analysis. The probes may be used as tools for surveying
CC tissues to detect the presence of expressed messages that contain their
CC specific exon, or in constructing genome-derived single exon microarrays.
CC In addition, the probes are used in identifying and characterising
CC alternative splicing events, in detecting and characterising gross
CC alterations in the genomic locus that includes their exon, in assessing
CC smaller genomic alterations, in priming the synthesis of nucleic acids,
CC or in expressing the ORF-encoded peptide. The present sequence is a human
CC single exon probe of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20030194704
XX
SQ Sequence 491 BP; 95 A; 163 C; 145 G; 88 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2.54 Length: 491
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 31.03% Indels: 0
DB: 12 Gaps: 0

US-10-624-218-1 (1-29) x ACH88921 (1-491)
QY 1 ArgAlaAlaPheLeuGlutylsGluAen 9
DB 95 CGGGCGGCTTCTCTGGAGAGGAGAAC 121

RESULT 12
ACH75221
ID ACH75221 standard; DNA; 546 BP.
XX
AC ACH75221;
XX
DT 29-JUL-2004 (first entry)
XX
DE Human genome derived single exon probe #8416.
XX
KW Human; probe; ss; gene expression; single exon probe; microarray;
KW alternative splicing event; genomic alteration.
XX
OS Homo sapiens.
XX
PN US2003194704-A1.
XX
PD 16-OCT-2003.
XX
PF 03-APR-2002; 2002US-00029386.
XX
PR 03-APR-2002; 2002US-00029386.
XX
PA (PENN/) PENN S G.
PA (RANK/) RANK D R.
PA (HANZ/) HANZEL D K.
XX
PI Penn SG, Rank DR, Hanzel DK;
XX
XX WPI; 2004-119264/12.
XX
PT New human genome-derived single exon nucleic acid probes useful for human
PT gene expression analysis, for identifying or characterizing alternative
PT splicing events, for assessing genomic alterations or as tools for
PT surveying tissues.
XX
PS Claim 15; SEQ ID NO 8416; 80pp; English.
XX
XX
```


AC ADP13309;
XX 26-AUG-2004 (first entry)
XX Renal cell carcinoma differentially expressed gene #45.
DE ds; diagnosis; non-blood disease; solid tumor; gene expression;
XX peripheral blood mononuclear cell; renal cell carcinoma; prostate cancer;
KW head/neck cancer; differential expression.
XX Homo sapiens.
OS WO2004048933-A2.
XX 10-JUN-2004.
XX 21-NOV-2003; 2003WO-US037481.
XX 21-NOV-2002; 2002US-0427982P.
PR 03-APR-2003; 2003US-0459782P.
XX (AMHP) WYETH.
PA (TWIN/) TWINE N C.
PA (BURC/) BURCZYNSKI M E.
PA (TREP/) TREPICCHIO W L.
PA (DORN/) DORNER A.
PA (STOV/) STOVER J A.
PA (SLON/) SLONI D K.
XX Twine NC, Burczynski ME, Trepicchio WL, Dorner A, Stover JA;
PI Sloni DK;
PI WPI; 2004-460799/43.
DR Diagnosing non-blood disease such as solid tumor, involves comparing
XX differential expression profile of specific genes in peripheral blood
PT sample of subject with reference expression profile of specific genes.
XX Disclosure; SEQ ID NO 45; 350pp; English.
XX The invention relate to a method of diagnosing (M1) non-blood disease
CC such as solid tumor by providing peripheral blood sample of human having
CC non-blood disease, and comparing an expression profile of specific genes
CC in the peripheral blood sample to reference expression profile of the
CC genes, where each of the genes is differentially expressed in peripheral
CC blood mononuclear cells (PBMCs) of patients having the disease as
CC compared to PBMCs of normal humans. The method is useful for diagnosing
CC non-blood disease such as solid tumor. The solid tumor is chosen from
CC renal cell carcinoma (RCC), prostate cancer and head/neck cancer. The
CC peripheral blood sample comprises enriched PBMCs. The peripheral blood
CC sample is a whole blood sample (claimed). (M1) is useful for identifying
CC genes that are differentially expressed in peripheral blood samples
CC isolated at different stages of progression, development or treatment of
CC RCC and/or other solid tumors. This sequence corresponds to a gene that
CC is differentially expressed and detected by the method of the invention.
CC (Note: this sequence is not given as part of the printed specification
CC but was obtained from WIPO in electronic format at
CC ftp.wipo./pub/published_pct_sequences/).
XX SQ Sequence 1385 BP; 279 A; 445 C; 404 G; 254 T; 0 U; 3 Other;
Alignment Scores: 6.48 Length: 1385
Pred. No.: 9.00 Matches: 9
Score: 100.00% Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 31.03% Gaps: 0
DB: 12
US-10-624-218-1 (1-29) x ADP13309 (1-1385)
QY 1 ArgAlaAlaPheLeuGluLysGluAen 9
|||||

Db 255 CGGGCGGCTTCTCGAGAGGAAC 281
RESULT 15
ADRS2783
ID ADRS2783 standard; DNA; 1385 BP.
XX ADRS2783;
AC ADRS2783;
XX 18-NOV-2004 (first entry)
DT Drug therapy altered expressed gene #134.
XX drug activity monitoring; expression profile; gene expression;
XX peripheral blood sample; peripheral blood mononuclear cell; drug therapy;
KW CCI-779; immunosuppressant; rapamycin; mammalian target of rapamycin;
KW mTOR; ds.
XX Homo sapiens.
OS WO2004072265-A2.
XX 26-AUG-2004.
XX 11-FEB-2004; 2004WO-US004118.
PF 11-FEB-2003; 2003US-0446133P.
PR 03-APR-2003; 2003US-0459782P.
PR 23-JAN-2004; 2004US-0538246P.
XX (AMHP) WYETH.
PA (BURC/) BURCZYNSKI M.
PA (TWIN/) TWINE N.
PA (DORN/) DORNER A J.
PA (TREP/) TREPICCHIO W L.
XX Burczynski M, Twine N, Dorner AJ, Trepicchio WL;
PI WPI; 2004-642301/62.
XX Monitoring drug activities in vivo comprises comparing an expression
XX profile of a gene in a peripheral blood sample of a patient before and
XX after drug therapy.
XX Disclosure; SEQ ID NO 134; 136pp; English.
XX The invention relates to a method of monitoring drug activities in vivo
CC by comparing an expression profile of at least one gene in a peripheral
CC blood sample of a patient to a reference expression profile of the at
CC least one gene, where the at least one gene is differentially expressed
CC in peripheral blood mononuclear cells (PBMCs) of patients who have a non-
CC blood disease and are subjected to a drug therapy as compared to PBMCs
CC isolated from the patient before the drug therapy, and where the patient
CC has the non-blood disease and is being treated by the drug therapy. The
CC method, kit, and nucleic acid array are useful for monitoring drug
CC activities in vivo. The drug is especially CCI-779, an ester analogue of
CC the immunosuppressant rapamycin which is a potent inhibitor of the
CC mammalian target of rapamycin (mTOR). This sequence represents a gene
CC expressed in PBMC altered by the drug therapy. (Note: this sequence does
CC no form part of the printed specification but was obtained in electronic
CC format from WIPO at ftp.wipo.int/pub/published_pct_sequences/).
XX SQ Sequence 1385 BP; 279 A; 445 C; 404 G; 254 T; 0 U; 3 Other;
Alignment Scores: 6.48 Length: 1385
Pred. No.: 9.00 Matches: 9
Score: 100.00% Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 31.03% Gaps: 0
DB: 13
US-10-624-218-1 (1-29) x ADRS2783 (1-1385)

Qy 1 ArgAlaAlaPheLeuGluLysGluAsn 9
|||
Db 255 CGGCGGCGCTTCCTGGAGAGGAGAAC 281
|||

Search completed: July 27, 2005, 12:14:26
Job time : 240.403 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 27, 2005, 11:44:30 ; Search time 67.6667 Seconds
(without alignments)
701.261 Million cell updates/sec

Title: US-10-624-218-1

Perfect score: 29

Sequence: 1 RAAPFLEXENTALTREVAELEXEVCRCENI 29

Scoring table:

OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Word size: 1

Total number of hits satisfying chosen parameters: 2398959

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO.spool_p/US10624218/runat_26072005_121436_3711/app_query.fasta_1.796
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=olip2n.rn1 -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -SPART=1 -END=-1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10624218 @CEN 1 1 213 @runat_26072005_121436_3711 -NCPU=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THRAIDS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA:
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PTCUTS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	19	65.5	1165	4	US-09-949-016-5024
2	19	65.5	32584	4	US-09-949-016-16766
3	9	31.0	1370	4	US-09-949-016-5339
4	9	31.0	10550	4	US-09-949-016-17081
5	9	31.0	18228	4	US-09-949-016-11964
6	9	31.0	18229	4	US-09-949-016-14056
7	8	27.6	25	4	US-09-396-196G-96363
8	8	27.6	25	4	US-09-396-196G-96364
9	8	27.6	601	4	US-09-949-016-40823
10	8	27.6	1603	3	US-09-103-664A-3
11	8	27.6	1889	3	US-09-103-664A-10
12	8	27.6	2091	3	US-09-103-664A-1

ALIGNMENTS

RESULT 1

US-09-949-016-5024
; Sequence 5024, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5024
; LENGTH: 1165
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-5024

Alignment Scores:
Pred. No.: 2.53e-11 Length: 1165
Score: 19.00 Matches: 19
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 65.52% Indels: 0
DB: 4 Gaps: 0

US-10-624-218-1 (1-29) x US-09-949-016-5024 (1-1165)

QY 1 ArgAlaAlaPheLeuGluLysGluAsnThrAlaLeuArgThrGluValaGluLeu 19

Sequence 112, App
Sequence 12900, A
Sequence 12412, A
Sequence 96362, A
Sequence 21922, A
Sequence 9464, Ap
Sequence 257, App
Sequence 88689, A
Sequence 93464, A
Sequence 94975, A
Sequence 94976, A
Sequence 95087, A
Sequence 95088, A
Sequence 177751, A
Sequence 181702, A
Sequence 181703, A
Sequence 181821, A
Sequence 181822, A
Sequence 181940, A
Sequence 181941, A
Sequence 1288, Ap
Sequence 3, Appli
Sequence 1947, Ap
Sequence 6808, Ap
Sequence 563, App
Sequence 17384, A
Sequence 1077, Ap
Sequence 16861, A
Sequence 9, Appli
Sequence 16791, A
Sequence 1, Appli
Sequence 25, Appli
Sequence 13420, A

DB: 4 Gaps: 0

US-10-624-218-1 (1-29) x US-09-949-016-5339 (1-1370)

Qy 1 ArgAlaAlaPheLeuGluLysGluAsn 9

Db 974 CGGGCGGCTTCTCTGGAGAGGAGAAC 1000

RESULT 4

US-09-949-016-17081

; Sequence 17081, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; PRIOR FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 17081

; LENGTH: 10550

; TYPE: DNA

; ORGANISM: Human

US-09-949-016-17081

Alignment Scores:

Pred. No.: 7.47 Length: 10550

Score: 9.00 Matches: 9

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 31.03% Indels: 0

DB: 4 Gaps: 0

US-10-624-218-1 (1-29) x US-09-949-016-17081 (1-10550)

Qy 1 ArgAlaAlaPheLeuGluLysGluAsn 9

Db 8154 CGGGCGGCTTCTCTGGAGAGGAGAAC 8180

RESULT 5

US-09-949-016-11964/c

; Sequence 11964, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; PRIOR FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 11964

; LENGTH: 18228

; TYPE: DNA

; ORGANISM: Human

US-09-949-016-11964

Alignment Scores:

DB: 991 CGGGCGGCTTCTCTGGAGAGGAGAACACAGCCCTGCGGAGGTGGCCGAGCTA 1047

RESULT 2

US-09-949-016-16766

; Sequence 16766, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; PRIOR FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 16766

; LENGTH: 32584

; TYPE: DNA

; ORGANISM: Human

US-09-949-016-16766

Alignment Scores:

Pred. No.: 5.29e-10 Length: 32584

Score: 19.00 Matches: 19

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 65.52% Indels: 0

DB: 4 Gaps: 0

US-10-624-218-1 (1-29) x US-09-949-016-16766 (1-32584)

Qy 1 ArgAlaAlaPheLeuGluLysGluAsnThrAlaLeuArgThrGluValAlaGluLeu 19

Db 30410 CGGGCGGCTTCTCTGGAGAGGAGAACACAGCCCTGCGGAGGTGGCCGAGCTA 30466

RESULT 3

US-09-949-016-5339

; Sequence 5339, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; PRIOR FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 5339

; LENGTH: 1370

; TYPE: DNA

; ORGANISM: Human

US-09-949-016-5339

Alignment Scores:

Pred. No.: 1.16 Length: 1370

Score: 9.00 Matches: 9

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 31.03% Indels: 0

Pred. No.: 12.3 Length: 18228
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 31.03% Indels: 0
DB: 4 Gaps: 0

US-10-624-218-1 (1-29) x US-09-949-016-11964 (1-18228)

Qy 17 AlaGluLeuGluLysGluValGlyArg 25
|||||
Db 10287 GCTGAGCTGGAGAAAGAGGTTTGGAGA 10261

RESULT 6
US-09-949-016-14056/c
; Sequence 14056, Application US/09949016
; Patent No. 6812139
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14056
; LENGTH: 18229
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-14056

Alignment Scores:
Pred. No.: 12.3 Length: 18229
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 31.03% Indels: 0
DB: 4 Gaps: 0

US-10-624-218-1 (1-29) x US-09-949-016-14056 (1-18229)

Qy 17 AlaGluLeuGluLysGluValGlyArg 25
|||||
Db 10287 GCTGAGCTGGAGAAAGAGGTTTGGAGA 10261

RESULT 7
US-09-396-196G-96363
; Sequence 96363, Application US/09396196G
; Patent No. 6821724
; GENERAL INFORMATION:
; APPLICANT: Michael Mittmann
; APPLICANT: David Mack
; APPLICANT: David Lockhart
; APPLICANT: Affymetrix, Inc.
; TITLE OF INVENTION: METHODS OF GENETIC ANALYSIS
; FILE REFERENCE: 3101.1
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: 60/100,678
; PRIOR FILING DATE: 1998-09-17
; NUMBER OF SEQ ID NOS: 127806
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 96363
; LENGTH: 25
; TYPE: DNA
; ORGANISM: mus musculus

US-09-396-196G-96363

Alignment Scores:
Pred. No.: 0.346 Length: 25
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 27.59% Indels: 0
DB: 4 Gaps: 0

US-10-624-218-1 (1-29) x US-09-396-196G-96363 (1-25)

Qy 3 AlaPheLeuGluLysGluAsnThr 10
|||||
Db 2 CGGTTCTCGGAGAGGAGACACA 25

RESULT 8
US-09-396-196G-96364
; Sequence 96364, Application US/09396196G
; Patent No. 6821724
; GENERAL INFORMATION:
; APPLICANT: Michael Mittmann
; APPLICANT: David Mack
; APPLICANT: David Lockhart
; APPLICANT: Affymetrix, Inc.
; TITLE OF INVENTION: METHODS OF GENETIC ANALYSIS
; FILE REFERENCE: 3101.1
; CURRENT APPLICATION NUMBER: US/09/396,196G
; CURRENT FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: 60/100,678
; PRIOR FILING DATE: 1998-09-17
; NUMBER OF SEQ ID NOS: 127806
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 96364
; LENGTH: 25
; TYPE: DNA
; ORGANISM: mus musculus
US-09-396-196G-96364

Alignment Scores:
Pred. No.: 0.346 Length: 25
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 27.59% Indels: 0
DB: 4 Gaps: 0

US-10-624-218-1 (1-29) x US-09-396-196G-96364 (1-25)

Qy 12 LeuArgThrGluValAlaGluLeu 19
|||||
Db 1 CTGCGAGCGAGGTTTCCGAGCTT 24

RESULT 9
US-09-949-016-40823
; Sequence 40823, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 40823

```
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-40823

Alignment Scores:
Pred. No.: 6.28 Length: 601
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 27.59% Indels: 0
DB: 4 Gaps: 0

US-10-624-218-1 (1-29) x US-09-949-016-40823 (1-601)

Qy 4 PhleuGluLysGluuAsnThrAla 11
Db 32 TTTTGAAGAAAGAAACACTGCT 55

RESULT 10
US-09-103-664A-3
; Sequence 3, Application US/09103664A
; Patent No. 6458358
; GENERAL INFORMATION:
; APPLICANT: University of Kentucky Research Foundation
; APPLICANT: Timoney, John
; APPLICANT: Artiushin, Sergey
; TITLE OF INVENTION: Compounds Encoding the Protective M-Like Protein of Streptococcus
; FILE REFERENCE: 50229-212
; CURRENT APPLICATION NUMBER: US/09/103,664A
; CURRENT FILING DATE: 1998-06-23
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 1603
; TYPE: DNA
; ORGANISM: Streptococcus equi
US-09-103-664A-3

Alignment Scores:
Pred. No.: 15.4 Length: 1603
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 27.59% Indels: 0
DB: 3 Gaps: 0

US-10-624-218-1 (1-29) x US-09-103-664A-3 (1-1603)

Qy 16 ValAlaGluLeuGluLysGluVal 23
Db 1171 GTAGCAGAGCTTGAAAAAGAGTT 1194

RESULT 11
US-09-103-664A-10
; Sequence 10, Application US/09103664A
; Patent No. 6458358
; GENERAL INFORMATION:
; APPLICANT: University of Kentucky Research Foundation
; APPLICANT: Timoney, John
; APPLICANT: Artiushin, Sergey
; TITLE OF INVENTION: Compounds Encoding the Protective M-Like Protein of Streptococcus
; FILE REFERENCE: 50229-212
; CURRENT APPLICATION NUMBER: US/09/103,664A
; CURRENT FILING DATE: 1998-06-23
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 10
; LENGTH: 1889
; TYPE: DNA
; ORGANISM: Streptococcus equi
```

```
US-09-103-664A-10

Alignment Scores:
Pred. No.: 17.9 Length: 1889
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 27.59% Indels: 0
DB: 3 Gaps: 0

US-10-624-218-1 (1-29) x US-09-103-664A-10 (1-1889)

Qy 16 ValAlaGluLeuGluLysGluVal 23
Db 1309 GTAGCAGAGCTTGAAAAAGAGTT 1332

RESULT 12
US-09-103-664A-1
; Sequence 1, Application US/09103664A
; Patent No. 6458358
; GENERAL INFORMATION:
; APPLICANT: University of Kentucky Research Foundation
; APPLICANT: Timoney, John
; APPLICANT: Artiushin, Sergey
; TITLE OF INVENTION: Compounds Encoding the Protective M-Like Protein of Streptococcus
; FILE REFERENCE: 50229-212
; CURRENT APPLICATION NUMBER: US/09/103,664A
; CURRENT FILING DATE: 1998-06-23
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 2091
; TYPE: DNA
; ORGANISM: Streptococcus equi
US-09-103-664A-1

Alignment Scores:
Pred. No.: 19.6 Length: 2091
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 27.59% Indels: 0
DB: 3 Gaps: 0

US-10-624-218-1 (1-29) x US-09-103-664A-1 (1-2091)

Qy 16 ValAlaGluLeuGluLysGluVal 23
Db 1511 GTAGCAGAGCTTGAAAAAGAGTT 1534

RESULT 13
US-09-566-921-112
; Sequence 112, Application US/09566921
; Patent No. 6682888
; GENERAL INFORMATION:
; APPLICANT: Loring, Jeanne F.
; APPLICANT: Tingley, Debora W.
; APPLICANT: Edwards, Carla M.
; TITLE OF INVENTION: GENES EXPRESSED IN ALZHEIMER'S DISEASE
; FILE REFERENCE: FA-0024 US
; CURRENT APPLICATION NUMBER: US/09/566,921
; CURRENT FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 138
; SOFTWARE: PERL Program
; SEQ ID NO 112
; LENGTH: 5773
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6682888 109082.18
US-09-566-921-112
```

Alignment Scores:
Pred. No.: 49.5 Length: 5773
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 27.59% Indels: 0
DB: 4 Gaps: 0

US-10-624-218-1 (1-29) x US-09-566-921-112 (1-5773)

Qy 17 AlaGluLeuGluLysGluValGly 24
Db 2323 GCAGAACTGGAGAGGAGGTGGC 2346

RESULT 14

US-09-949-016-12900/c
; Sequence 12900, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12900
; LENGTH: 462589
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12900

Alignment Scores:
Pred. No.: 2.7e+03 Length: 462589
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 27.59% Indels: 0
DB: 4 Gaps: 0

US-10-624-218-1 (1-29) x US-09-949-016-12900 (1-462589)

Qy 4 PheLeuGluLysGluAsnThrAla 11
Db 281446 TTTTGAAGAAAAGAAACACTGCT 281423

RESULT 15

US-09-949-016-12412/c
; Sequence 12412, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 12412
; LENGTH: 476044
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12412
Alignment Scores:
Pred. No.: 2.77e+03 Length: 476044
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 27.59% Indels: 0
DB: 4 Gaps: 0

US-10-624-218-1 (1-29) x US-09-949-016-12412 (1-476044)

Qy 4 PheLeuGluLysGluAsnThrAla 11
Db 294902 TTTTGAAGAAAAGAAACACTGCT 294879

Search completed: July 27, 2005, 16:47:55
Job time : 133.667 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 27, 2005, 11:55:09 ; Search time 368.34 Seconds
(without alignments)
509.100 Million cell updates/sec

Title: US-10-624-218-1

Perfect score: 29

Sequence: 1 RAAFLXKENTALTREVAEKEVGRCENI 29

Scoring table:

Oligo 60.0, Xgapext 60.0
Ygapop 60.0, Ygapext 60.0
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 7277826 seqs, 3233139505 residues

Word size: 1

Total number of hits satisfying chosen parameters: 14543013

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters: -MODEL=frame+ p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO.spool.p/US10624218/runat_26072005_121437_3743/app_query.fasta_1.796
-DB=Published Applications NA -QWITS=fastap -SUFFIX=olip2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=quality -THR MIN=1
-ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US10624218@cgn2_1.1189@runat_26072005_121437_3743
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60
-FCAPOP=6 -FCAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : Published Applications NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:
19: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq:
20: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq:
21: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq:
22: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:
23: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq:
24: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:
25: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:
26: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	9	31.0	217	16	US-10-029-386-24227
C 2	9	31.0	466	10	US-09-918-995-23172
C 3	9	31.0	491	16	US-10-029-386-22116
C 4	9	31.0	546	16	US-10-029-386-8416
C 5	9	31.0	546	15	US-10-029-386-10527
C 6	9	31.0	1385	19	US-10-717-597-45
C 7	9	31.0	1385	19	US-10-775-169-134
C 8	9	31.0	1478	21	US-10-887-553A-202
C 9	9	31.0	1638	21	US-10-764-420-2185
C 10	9	31.0	1671	17	US-10-388-934-173
C 11	9	31.0	1671	21	US-10-870-387-20
C 12	8	27.6	25	21	US-10-719-900-429122
C 13	8	27.6	25	21	US-10-719-900-429125
C 14	8	27.6	25	21	US-10-809-189-96363
C 15	8	27.6	25	21	US-10-809-189-96364
C 16	8	27.6	466	19	US-10-437-963-75775
C 17	8	27.6	516	19	US-10-021-323-13045
C 18	8	27.6	546	19	US-10-021-323-16850
C 19	8	27.6	571	19	US-10-021-323-15000
C 20	8	27.6	600	22	US-10-972-079-81913
C 21	8	27.6	651	18	US-10-424-599-30989
C 22	8	27.6	1701	17	US-10-282-122A-35417
C 23	8	27.6	5432	9	US-09-962-832-140
C 24	8	27.6	5432	9	US-09-954-456-309
C 25	8	27.6	5432	10	US-09-960-706-462
C 26	8	27.6	5432	10	US-09-873-319-282
C 27	8	27.6	5432	21	US-10-843-641A-3336
C 28	8	27.6	5432	21	US-10-843-641A-6026
C 29	8	27.6	5773	22	US-10-765-700-112
C 30	7	24.1	199	9	US-09-809-189-96362
C 31	7	24.1	199	9	US-09-867-701-2609
C 32	7	24.1	202	19	US-10-674-124A-3492
C 33	7	24.1	211	20	US-10-425-115-130734
C 34	7	24.1	257	9	US-09-867-701-2637
C 35	7	24.1	326	17	US-10-012-697-876
C 36	7	24.1	330	9	US-09-880-107-893
C 37	7	24.1	381	11	US-09-864-408A-2041
C 38	7	24.1	392	20	US-10-425-115-77176
C 39	7	24.1	409	20	US-10-425-115-78684
C 40	7	24.1	440	9	US-09-867-701-10557
C 41	7	24.1	492	17	US-10-320-797-2020
C 42	7	24.1	536	22	US-10-972-079-93228
C 43	7	24.1	542	19	US-10-021-323-15743
C 44	7	24.1	584	16	US-10-029-386-6891
C 45	7	24.1	589	18	US-10-425-114-23262

ALIGNMENTS

RESULT 1
US-10-029-386-24227/c
; Sequence 24227, Application US/10029386
; Publication NO. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 24227

```

; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 22116
; LENGTH: 491
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR19.3
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.7
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.79
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.3
; OTHER INFORMATION: SWISSPROT HIT: Q10586, EVALUE 4.00e+00
; OTHER INFORMATION: NT HIT: D28468.1, EVALUE 0.00e+00
; OTHER INFORMATION: EST_HUMAN HIT: A1684518.1, EVALUE 0.00e+00
US-10-029-386-22116

Alignment Scores:
Pred. No.:      2.18      Length:      491
Score:          9.00      Matches:       9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match:    31.03% Indels:         0
DB:             16      Gaps:         0

US-10-624-218-1 (1-29) x US-10-029-386-22116 (1-491)
QY   1 ArgAlaAlaPheLeuGluLysGluAsn 9
Db   95 CGGGCGGCCTTCCTGGAGGAAGAAC 121

RESULT 4
US-10-029-386-8416
; Sequence 8416, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR C
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 8416
; LENGTH: 546
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR19.3
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.7
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.79
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.3
; OTHER INFORMATION: SWISSPROT HIT: Q10586, EVALUE 4.00e+00
; OTHER INFORMATION: NT HIT: U48213.1, EVALUE 0.00e+00
; OTHER INFORMATION: EST_HUMAN HIT: A1684518.1, EVALUE 0.00e+00
US-10-029-386-8416

Alignment Scores:
Pred. No.:      2.39      Length:      546
Score:          9.00      Matches:       9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
;

```



```
Query Match: 31.03% Indels: 0
DB: 16 Gaps: 0

US-10-624-218-1 (1-29) x US-10-029-386-8416 (1-546)

QY 1 ArgAlaAlaPheLeuGluLysGluAsn 9
DB 120 CGGGCGGCCTTCTCGAGAAGGAGAAC 146

RESULT 5
US-10-029-386-10527/c
; Sequence 10527, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR C
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AROMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 10527
; LENGTH: 546
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR19.3
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.6
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.4
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 8.7
; OTHER INFORMATION: NT HIT: U48213.1, EVALUE 0.00e+00
; OTHER INFORMATION: EST_HUMAN HIT: A1684518.1, EVALUE 0.00e+00
; OTHER INFORMATION: SWISSPROT HIT: Q10586, EVALUE 4.00e-34
US-10-029-386-10527

Alignment Scores:
Pred. No.: 2.39 Length: 546
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 31.03% Indels: 0
DB: 16 Gaps: 0

US-10-624-218-1 (1-29) x US-10-029-386-10527 (1-546)

QY 1 ArgAlaAlaPheLeuGluLysGluAsn 9
DB 427 CGGGCGGCCTTCTCGAGAAGGAGAAC 401

RESULT 6
US-10-717-597-45
; Sequence 45, Application US/10717597
; Publication No. US20040110221A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Burczynski, Michael E.
; APPLICANT: Twine, Natalie C.
; APPLICANT: Dörner, Andrew J.
; APPLICANT: Trepicchio, William L.
; APPLICANT: Slonim, Donna K.
; APPLICANT: Stover, Jennifer A.
; TITLE OF INVENTION: METHODS FOR DIAGNOSING RCC AND OTHER SOLID TUMORS
; FILE REFERENCE: AM101080
; CURRENT APPLICATION NUMBER: US/10/717,597
; CURRENT FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: US 60/459,782
; PRIOR FILING DATE: 2003-04-03
; PRIOR APPLICATION NUMBER: US 60/427,982
; PRIOR FILING DATE: 2002-11-21
```

```
; NUMBER OF SEQ ID NOS: 4904
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 45
; LENGTH: 1385
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (20)..(1158)
; OTHER INFORMATION: The "n" residue at position 1158 is nucleotide g according to th
; OTHER INFORMATION: corresponding sequence (DBP gene) in the Entrez Human Genome
; OTHER INFORMATION: Sequence Database. No residue exists at n position 20 and 25
; OTHER INFORMATION: according to the same corresponding sequence.
US-10-717-597-45

Alignment Scores:
Pred. No.: 5.32 Length: 1385
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 31.03% Indels: 0
DB: 19 Gaps: 0

US-10-624-218-1 (1-29) x US-10-717-597-45 (1-1385)

QY 1 ArgAlaAlaPheLeuGluLysGluAsn 9
DB 255 CGGGCGGCCTTCTCGAGAAGGAGAAC 281

RESULT 7
US-10-775-169-134
; Sequence 134, Application US/10775169
; Publication No. US20040175743A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Burczynski, Michael
; APPLICANT: Twine, Natalie
; APPLICANT: Dörner, Andrew
; APPLICANT: Trepicchio, William
; TITLE OF INVENTION: Method for Monitoring Drug Activities In Vivo
; FILE REFERENCE: AM101080 (031896-013000)
; CURRENT APPLICATION NUMBER: US/10/775,169
; CURRENT FILING DATE: 2004-02-11
; NUMBER OF SEQ ID NOS: 5278
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 134
; LENGTH: 1385
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (20)..(1158)
; OTHER INFORMATION: Each "n" represents a nucleotide selected from a, t, g or c, or
; OTHER INFORMATION: contains no nucleotide.
US-10-775-169-134

Alignment Scores:
Pred. No.: 5.32 Length: 1385
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 31.03% Indels: 0
DB: 19 Gaps: 0

US-10-624-218-1 (1-29) x US-10-775-169-134 (1-1385)

QY 1 ArgAlaAlaPheLeuGluLysGluAsn 9
DB 255 CGGGCGGCCTTCTCGAGAAGGAGAAC 281

RESULT 8
US-10-887-553A-202
; Sequence 202, Application US/10887553A
```

```

; Publication No. US20050085436A1
; GENERAL INFORMATION:
; APPLICANT: Garza, Dan
; APPLICANT: Li, Hao
; TITLE OF INVENTION: Method to treat conditions associated
; TITLE OF INVENTION: with insulin signalling dysregulation
; FILE REFERENCE: 4-33262
; CURRENT APPLICATION NUMBER: US/10/887,553A
; CURRENT FILING DATE: 2004-07-08
; PRIOR APPLICATION NUMBER: 60/485,883
; PRIOR FILING DATE: 2003-08-07
; NUMBER OF SEQ ID NOS: 1208
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 202
; LENGTH: 1478
; TYPE: DNA
; ORGANISM: human
US-10-887-553A-202

Alignment Scores:
Pred. No.: 5.63 Length: 1478
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 31.03% Indels: 0
Gaps: 0
DB:

US-10-624-218-1 (1-29) x US-10-887-553A-202 (1-1478)

QY 1 ArgAlaAlaPheLeuGluLysGluAsn 9
DB 1193 CGGGCGGCTTCTCTGGAGAGGAAAC 1219

RESULT 9
; Sequence 2185, Application US/10764420
; Publication No. US20050084872A1
; GENERAL INFORMATION:
; APPLICANT: Lum, Pek Yee
; APPLICANT: Tan, Yejun
; APPLICANT: Dai, Hongyue
; TITLE OF INVENTION: Methods For Determining Whether An Agent
; TITLE OF INVENTION: Possesses A Defined Biological Activity
; FILE REFERENCE: ROSA122057
; CURRENT APPLICATION NUMBER: US/10/764,420
; CURRENT FILING DATE: 2004-01-23
; PRIOR APPLICATION NUMBER: US 60/442,797
; PRIOR FILING DATE: 2003-01-24
; PRIOR APPLICATION NUMBER: US 60/474,413
; PRIOR FILING DATE: 2003-05-30
; NUMBER OF SEQ ID NOS: 3683
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2185
; LENGTH: 1638
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-764-420-2185

Alignment Scores:
Pred. No.: 6.15 Length: 1638
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 31.03% Indels: 0
Gaps: 0
DB:

US-10-624-218-1 (1-29) x US-10-764-420-2185 (1-1638)

QY 1 ArgAlaAlaPheLeuGluLysGluAsn 9
DB 1246 AGGGCTGCTTCTCTGGAGAGGAAAC 1272

RESULT 10
```

```

US-10-388-934-173
; Sequence 173, Application US/10388934
; Publication No. US20040005547A1
; GENERAL INFORMATION:
; APPLICANT: Boess, Franziska
; APPLICANT: Suter-Dick, Laura
; APPLICANT: Wolf, Detlef
; TITLE OF INVENTION: BIOMARKERS AND EXPRESSION PROFILES FOR TOXICOLOGY
; FILE REFERENCE: 21199
; CURRENT APPLICATION NUMBER: US/10/388,934
; CURRENT FILING DATE: 2003-03-14
; PRIOR APPLICATION NUMBER: 02005336.9
; PRIOR FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 02015657.6
; PRIOR FILING DATE: 2002-07-17
; NUMBER OF SEQ ID NOS: 862
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 173
; LENGTH: 1671
; TYPE: DNA
; ORGANISM: Rattus norvegicus (No. US20040005547Alway rat)
US-10-388-934-173

Alignment Scores:
Pred. No.: 6.25 Length: 1671
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 31.03% Indels: 0
Gaps: 0
DB:

US-10-624-218-1 (1-29) x US-10-388-934-173 (1-1671)

QY 1 ArgAlaAlaPheLeuGluLysGluAsn 9
DB 1223 CGGGCAGCCTTCTCTGGAGAGGAAAC 1249

RESULT 11
US-10-870-387-20
; Sequence 20, Application US/10870387
; Publication No. US20050064462A1
; GENERAL INFORMATION:
; APPLICANT: Stein, Bernd
; APPLICANT: Zhu, Dan
; APPLICANT: Patnaik, Nandita
; APPLICANT: Wang, Yuedi
; TITLE OF INVENTION: METHODS, COMPOSITIONS, AND KITS FOR PREDICTING THE EFFECT OF COM
; TITLE OF INVENTION: HOT FLASH SYMPTOMS
; FILE REFERENCE: 10624-127-888
; CURRENT APPLICATION NUMBER: US/10/870,387
; CURRENT FILING DATE: 2004-06-16
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 20
; LENGTH: 1671
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-10-870-387-20

Alignment Scores:
Pred. No.: 6.25 Length: 1671
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 31.03% Indels: 0
Gaps: 0
DB:

US-10-624-218-1 (1-29) x US-10-870-387-20 (1-1671)

QY 1 ArgAlaAlaPheLeuGluLysGluAsn 9
DB 1223 CGGGCAGCCTTCTCTGGAGAGGAAAC 1249
```

```
RESULT 12
US-10-719-900-429122
; Sequence 429122, Application US/10719900
; Publication No. US20050026164A1
; GENERAL INFORMATION:
; APPLICANT: Xue Mei Zhou
; TITLE OF INVENTION: Methods of Genetic Analysis of Mouse
; FILE REFERENCE: 3528.1
; CURRENT APPLICATION NUMBER: US/10/719,900
; PRIOR FILING DATE: 2003-11-20
; PRIOR APPLICATION NUMBER: 60/427,808
; PRIOR FILING DATE: 2002 11 20
; NUMBER OF SEQ ID NOS: 982914
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 429122
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-719-900-429122

Alignment Scores:
Pred. No.: 1.93 Length: 25
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 27.59% Indels: 0
DB: 21 Gaps: 0

US-10-624-218-1 (1-29) x US-10-719-900-429122 (1-25)

QY 7 LysGluAsnThrAlaLeuArgThr 14
|||||
DB 2 AAGGAGAACACAGCCCTGGGACG 25

RESULT 13
US-10-719-900-429125
; Sequence 429125, Application US/10719900
; Publication No. US20050026164A1
; GENERAL INFORMATION:
; APPLICANT: Xue Mei Zhou
; TITLE OF INVENTION: Methods of Genetic Analysis of Mouse
; FILE REFERENCE: 3528.1
; CURRENT APPLICATION NUMBER: US/10/719,900
; PRIOR FILING DATE: 2003-11-20
; PRIOR APPLICATION NUMBER: 60/427,808
; PRIOR FILING DATE: 2002 11 20
; NUMBER OF SEQ ID NOS: 982914
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 429125
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-719-900-429125

Alignment Scores:
Pred. No.: 1.93 Length: 25
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 27.59% Indels: 0
DB: 21 Gaps: 0

US-10-624-218-1 (1-29) x US-10-719-900-429125 (1-25)

QY 7 LysGluAsnThrAlaLeuArgThr 14
|||||
DB 2 AAGGAGAACACTGCCCTGGGACG 25

RESULT 14
US-10-809-189-96363
; Sequence 96363, Application US/10809189
; Publication No. US20050048531A1
; GENERAL INFORMATION:
; APPLICANT: Michael Mittmann
; APPLICANT: David Lockhart
; APPLICANT: Affymetrix, Inc.
; TITLE OF INVENTION: Methods of Genetic Analysis
; FILE REFERENCE: 3101.1
; CURRENT APPLICATION NUMBER: US/10/809,189
; CURRENT FILING DATE: 2004-03-25
; PRIOR APPLICATION NUMBER: US/09/396,196
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: 60/100,678
; PRIOR FILING DATE: 1998-09-17
; NUMBER OF SEQ ID NOS: 127806
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 96363
; LENGTH: 25
; TYPE: DNA
; ORGANISM: mus musculus
US-10-809-189-96363

Alignment Scores:
Pred. No.: 1.93 Length: 25
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 27.59% Indels: 0
DB: 21 Gaps: 0

US-10-624-218-1 (1-29) x US-10-809-189-96363 (1-25)

QY 12 LeuArgThrGluValAlaGluLeu 19
|||||
DB 1 CTGGGACGAGGTTGCCGAGCTT 24
```

```
; APPLICANT: Michael Mittmann
; APPLICANT: David Mack
; APPLICANT: David Lockhart
; APPLICANT: Affymetrix, Inc.
; TITLE OF INVENTION: Methods of Genetic Analysis
; FILE REFERENCE: 3101.1
; CURRENT APPLICATION NUMBER: US/10/809,189
; CURRENT FILING DATE: 2004-03-25
; PRIOR APPLICATION NUMBER: US/09/396,196
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: 60/100,678
; PRIOR FILING DATE: 1998-09-17
; NUMBER OF SEQ ID NOS: 127806
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 96363
; LENGTH: 25
; TYPE: DNA
; ORGANISM: mus musculus
US-10-809-189-96363

Alignment Scores:
Pred. No.: 1.93 Length: 25
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 27.59% Indels: 0
DB: 21 Gaps: 0

US-10-624-218-1 (1-29) x US-10-809-189-96363 (1-25)

QY 3 AlapheLeuGluLysGluAsnThr 10
|||||
DB 2 GCGTTCTGGAGAGGAGACACA 25

RESULT 15
US-10-809-189-96364
; Sequence 96364, Application US/10809189
; Publication No. US20050048531A1
; GENERAL INFORMATION:
; APPLICANT: Michael Mittmann
; APPLICANT: David Mack
; APPLICANT: David Lockhart
; APPLICANT: Affymetrix, Inc.
; TITLE OF INVENTION: Methods of Genetic Analysis
; FILE REFERENCE: 3101.1
; CURRENT APPLICATION NUMBER: US/10/809,189
; CURRENT FILING DATE: 2004-03-25
; PRIOR APPLICATION NUMBER: US/09/396,196
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: 60/100,678
; PRIOR FILING DATE: 1998-09-17
; NUMBER OF SEQ ID NOS: 127806
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 96364
; LENGTH: 25
; TYPE: DNA
; ORGANISM: mus musculus
US-10-809-189-96364

Alignment Scores:
Pred. No.: 1.93 Length: 25
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 27.59% Indels: 0
DB: 21 Gaps: 0

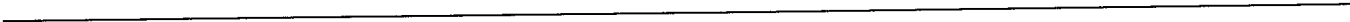
US-10-624-218-1 (1-29) x US-10-809-189-96364 (1-25)

QY 12 LeuArgThrGluValAlaGluLeu 19
|||||
DB 1 CTGGGACGAGGTTGCCGAGCTT 24
```

us-10-624-218-1.olip2n.rnpb

Thu Jul 28 11:34:40 2005

Search completed: July 27, 2005, 17:17:38
Job time : 372.34 secs



GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 27, 2005, 11:14:04 ; Search time 2016.31 Seconds

(without alignments)
547.469 Million cell updates/sec

Title: US-10-624-218-1

Perfect score: 29

Sequence: 1 RAAFLEKENTALTEVALEKEVGCRCNI 29

Scoring table:

OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Word size: 1

Total number of hits satisfying chosen parameters: 68476236

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame_plus_p2n.model -DEV=xlp
-O=/cgn2_1/USPTO_spoos_p/US10624218/runat_26072005_121436_3698/app.query.fasta_1.796
-DB=EST -QFMT=fastap -SUFFIX=olip2n.rst -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc
-NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10624218 @C@N 1 1 9525 @runat_26072005_121436_3698 -NCPU=6 -ICPU=3
-NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THRSADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gse1:*
9: gb_gse2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	19	65.5	341	1	AI014348 am51g11.s
2	19	65.5	351	6	BY775604 BY775604
3	19	65.5	380	1	AI892971 mg82h05.Y
4	19	65.5	398	6	CB707074 CB707074
5	19	65.5	418	7	CN394275 CN394275
6	19	65.5	440	2	BE295612 BE295612
7	19	65.5	446	1	AA138848 mg82h05.r
8	19	65.5	464	4	BG986696 BG986696
9	19	65.5	484	5	BU262464 BU262464

10	19	65.5	528	2	BF660216
11	19	65.5	540	7	CR536622
12	19	65.5	555	2	AW836791
13	19	65.5	557	1	BU111019
14	19	65.5	578	1	AL120537
15	19	65.5	599	2	AW500333
16	19	65.5	626	5	BU469699
17	19	65.5	631	6	CB577428
18	19	65.5	638	7	CK624751
19	19	65.5	654	4	BG432348
20	19	65.5	655	9	CR828685
21	19	65.5	673	5	BM963436
22	19	65.5	693	6	CB247229
23	19	65.5	786	5	BU466941
24	19	65.5	788	7	CN455515
25	19	65.5	812	5	BQ442190
26	19	65.5	850	2	BF309490
27	19	65.5	928	6	CB590113
28	19	65.5	933	5	BQ887766
29	19	65.5	941	2	BE737088
30	19	65.5	958	5	BU149688
31	16	55.2	563	7	CO794673
32	16	55.2	653	2	BF203280
33	15	51.7	498	2	AW836776
34	13	44.8	520	2	AW637693
35	13	44.8	591	4	BQ995767
36	13	44.8	694	5	EX852163
37	13	44.8	707	5	BP702661
38	13	44.8	709	5	EX846930
39	13	44.8	919	5	BU899556
40	13	44.8	922	5	BQ733819
41	12	41.4	619	2	BF306473
42	11	37.9	846	5	BU441005
43	10	34.5	1034	4	BM466864
44	9	31.0	268	7	F26608
45	9	31.0	277	2	BE720262

ALIGNMENTS

RESULT 1
AI014348
LOCUS
DEFINITION
am51g11.s1 Johnston frontal cortex Homo sapiens cDNA clone
IMAGE:1539140 3', similar to gb:M95586 TRANSCRIPTION FACTOR E2-ALPHA
(HUMAN); mRNA sequence.
AI014348 341 bp mRNA linear EST 15-JUN-1998
AI014348.1 GI:3228180
AI014348.1
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
REFERENCE
AUTHORS
Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S.,
Krisman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M.,
Martin, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F.,
Theising, B., White, Y., Wyllie, T., Waterston, R. and Wilson, R.
TITLE
JOURNAL
COMMENT
WashU-NCI human EST Project
Unpublished (1997)
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: esewatson.wustl.edu
This clone is available royalty-free through LML; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40m3 fwd. ET from Amersham
High quality sequence stop: 314.
Location/Qualifiers
1. .341
/organism="Homo sapiens"

```

/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1539140"
/sex="male"
/tissue_type="pooled frontal lobe"
/lab_host="adult"
/dev_stage="adult"
/clone_lib="Johnston frontal cortex"
/notes="Organ: brain; Vector: Bluescript SK-; Site_1:
ECORI; Stanley Neuropathology Consortium
(www.stanleylab.org) brains S-58, S-65, S-67, S-78.
Random + oligo-dT primed into EcoRI site of ZAP II Vector.
Mass excised. Avg insert length 1.9kb. Custom library
provided by Dr. Nancy Johnston [(410) 614-3918,
nlj@weichlink.welch.jhu.edu].

ORIGIN
Alignment Scores:
Pred. No.: 1.09e-09 Length: 341
Score: 19.00 Matches: 19
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 65.52% Indels: 0
DB: 1 Gaps: 0

US-10-624-218-1 (1-29) x AI014348 (1-341)

Qy 1 ArgAlaAlaPheLeuGluLysGluAnThrAlaLeuArgThrGluValAlaGluLeu 19
|||||
Db 131 CGGGCAGCCTTCTCGGAGAGGAGAACACAGCCCTGCGCAGGAGTGGCGAGCTA 187

RESULT 2
BY775604 351 bp mRNA linear EST 23-MAR-2004
LOCUS
DEFINITION
Mus musculus cDNA clone L930096D17 5', mRNA sequence.

ACCESSION
BY775604.1 GI:39702242
VERSION
KEYWORDS
SOURCE
Mus musculus (house mouse)
ORGANISM

REFERENCE
AUTHORS
Carninci, P., Waki, K., Shiraki, T., Konno, H., Shibata, K., Itoh, M.,
Aizawa, K., Arakawa, T., Ishii, Y., Sasaki, D., Bono, H., Kondo, S.,
Sugahara, Y., Saito, R., Osato, N., Fukuda, S., Sato, K., Watahiki, A.,
Hirozane-Kishikawa, T., Nakamura, M., Shibata, Y., Yasunishi, A.,
Kikuchi, N., Yoshiki, A., Kusakabe, M., Gustincich, S., Beisel, K.,
Pavan, W., Aidinis, V., Nakagawara, A., Heid, W. A., Iwata, H., Kono, T.,
Nakauchi, H., Lyons, P., Wells, C., Hume, D. A., Fagiolini, M.,
Henssch, T. K., Brinkmeier, M., Camper, S., Hirota, J., Mombaerts, P.,
Muramatsu, M., Okazaki, Y., Kawai, J., and Hayashizaki, Y.
Targeting a complex transcriptome: the construction of the mouse
full-length cDNA encyclopedia
Genome Res. 13 (6B), 1273-1289 (2003)
22703353
12819125
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gscc.riken.jp, URL: http://genome.gsc.riken.jp/
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site (http://genome.gsc.riken.jp/) for
further details.

```

```

Location/Qualifiers
1. 351
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="L930096D17"
/tissue_type="whole body"
/dev_stage="17.5 days embryo"
/clone_lib="RIKEN full-length enriched, 17.5 days embryo
whole body"

ORIGIN
Alignment Scores:
Pred. No.: 1.12e-09 Length: 351
Score: 19.00 Matches: 19
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 65.52% Indels: 0
DB: 6 Gaps: 0

US-10-624-218-1 (1-29) x BY775604 (1-351)

Qy 1 ArgAlaAlaPheLeuGluLysGluAnThrAlaLeuArgThrGluValAlaGluLeu 19
|||||
Db 24 CGGGCAGCCTTCTCGGAGAGGAGAACACAGCCCTGCGCAGGAGTGGCGAGCTT 80

RESULT 3
BY775604 380 bp mRNA linear EST 15-MAR-2000
LOCUS
DEFINITION
mg82h05.y1 Stratagene mouse melanoma (#937312) Mus musculus cDNA
clone IMAGE:585273 5', similar to gb:M95586 TRANSCRIPTION FACTOR
E2-ALPHA (HUMAN);, mRNA sequence.

ACCESSION
BY775604.1 GI:5598873
VERSION
KEYWORDS
SOURCE
Mus musculus (house mouse)
ORGANISM

REFERENCE
AUTHORS
Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T.,
Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y.,
Person, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurk, R.,
Ritter, E., Kohn, S., Shin, I., Jackson, Y., Cardenas, M., McCann, R.,
Waterston, R., and Wilson, R.
The WashU-NCI Mouse EST Project 1999
Unpublished (1999)
Contact: Marra M/WashU-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LML; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
This read is a RESEQUENCE of a previously sequenced mouse clone
This read has been verified (found to hit its original self in the
correct orientation)
MGI:359921
Seq primer: -40RP from Gibco
High quality sequence stop: 336.
Location/Qualifiers
1. 380
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:585273"
/tissue_type="melanoma"
/dev_stage="m2 cells"
/lab_host="SOLR (kanamycin resistant)"
/clone_lib="Stratagene mouse melanoma (#937312)"
/notes="Organ: skin; Vector: pBluescript SK-; Site_1:

```

EcokR; Site 2: XhoI; Cloned unidirectionally. Primer: Oligo dt. From M2 cells, a highly metastatic derivative of the K-1735 (mouse) melanoma. Average insert size: 1.0 kb; Uni-ZAP XR Vector: -5' adaptor sequence: 5' GAATTCGGCAGG 3' -3' adaptor sequence: 5' CTCAGTTTTTTTTTTTTTTT 3'

ORIGIN

Alignment Scores:

Pred. No.: 1.2e-09 Length: 380
Score: 19.00 Matches: 19
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 65.52% Indels: 0
DB: 1 Gaps: 0

US-10-624-218-1 (1-29) x AI892971 (1-380)

Qy 1 ArgAlaAlaPheLeuGluLysGluAsnThrAlaLeuArgThrGluValalaGluLeu 19
Db 89 CGGGCAGGCTTCTCGAGAGGAGAACACAGCCCTCGCGACGGAGGTTCGCGAGCTT 145

RESULT 4

CB707074 398 bp mRNA linear EST 10-APR-2003
LOCUS AMGNNUC:TRCP2-00001-D6-A trcp2 (10289) Rattus norvegicus cDNA clone
DEFINITION trcp2-00001-d6 5', mRNA sequence.

ACCESSION

CB707074

VERSION

CB707074.1 GI:29764222

KEYWORDS

EST.

SOURCE

Rattus norvegicus (Norway rat)

ORGANISM

Rattus norvegicus

REFERENCE

1 (bases 1 to 398)

Amgen EST Program.

Amgen Rat EST Program

Unpublished (2003)

Contact: Dan Fitzpatrick

Amgen, Inc

One Amgen Center Drive, Thousand Oaks, CA 91320-1799, USA

Tel: 805 447-4881

Plate: 00001 row: d column: 6.

Location/Qualifiers

1. 398

/organism="Rattus norvegicus"

/mol_type="mRNA"

/db_xref="taxon:10116"

/clone="trcp2-00001-d6"

/tissue type="choroid plexus brain"

/clone_lib="trcp2 (10289)"

/note="Vector: C6KFG7L; Site_1: SalI; Site_2: NotI;

choroid plexus brain region"

FEATURES

source

Alignment Scores:

Pred. No.: 1.26e-09 Length: 398
Score: 19.00 Matches: 19
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 65.52% Indels: 0
DB: 6 Gaps: 0

US-10-624-218-1 (1-29) x CB707074 (1-398)

Qy

1 ArgAlaAlaPheLeuGluLysGluAsnThrAlaLeuArgThrGluValalaGluLeu 19

Db

50 CGGGCAGGCTTCTCGAGAGGAGAACACAGCCCTCGCGACGGAGGTTCGCGAGCTT 106

RESULT 5

CN394275

LOCUS

DEFINITION

ACCESSION

CN394275

VERSION

CN394275.1 GI:47381870

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

REFERENCE

1 (bases 1 to 418)

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

Brandenberger, R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G.J.,

Li, Y., Xu, C., Fang, R., Guegler, K., Rao, M.S., Mandalam, R.,

Lebkowski, J. and Stanton, L.W.

Transcriptome characterization elucidates signaling networks that

control human ES cell growth and differentiation

Nat. Biotechnol. 22 (6), 707-716 (2004)

Contact: Brandenberger R

Regenerative Medicine

Genon Corporation

230 Constitution Drive, Menlo Park, CA 94025, USA

Tel: 650 473 8658

Fax: 650 473 7760

Email: rbrandenberger@genon.com

Insert Length: 418 Std Error: 0.00.

Location/Qualifiers

1. 418

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/tissue_type="embryonic stem cells, cell lines H1, H7, and H9"

/clone_lib="GRN ES"

/note="Oligo dT primed, full-length enriched cDNA library

from undifferentiated hES cell lines H1 (p32), H7 (p29),

and H9 (p26) maintained in feeder-free conditions"

US-10-624-218-1 (1-29) x CN394275 (1-418)

Qy

1 ArgAlaAlaPheLeuGluLysGluAsnThrAlaLeuArgThrGluValalaGluLeu 19

Db 68 CGGGCAGGCTTCTCGAGAGGAGAACACAGCCCTCGCGACGGAGGTTCGCGAGCTA 124

RESULT 6

BE295612

LOCUS

DEFINITION

601176577F1 NTH_MGC_17 Homo sapiens cDNA clone IMAGE:3531762 5',

mRNA sequence.

BE295612

BE295612.1 GI:9179163

EST.

Accession

Source

Organism

Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

NIH-MGC http://mgi.nci.nih.gov/.

1 (bases 1 to 440)

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgaabs-r@mail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Ling Hong/Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov

Plate: L1CM206 row: j column: 19
High quality sequence start: 6
High quality sequence stop: 437.
Location/Qualifiers
1. .440
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3531762"
/tissue type="rhabdomyosarcoma"
/lab host="DH10B (phage-resistant)"
/clone lib="NIH_MGC 17"
/note="Organ: muscle; Vector: pOTB7; Site 1: EcoRI; Site 2: XhoI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

Alignment Scores:
Pred. No.: 1.38e-09 Length: 440
Score: 19.00 Matches: 19
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 65.52% Indels: 0
Gaps: 2
DB:

US-10-624-218-1 (1-29) x BE295612 (1-440)

Qy 1 ArgAlaAlaPheLeuGluLysThrAlaLeuArgThrGluValAlaGluLeu 19
|||||

Db 327 CGGCAGCCTTCCTGGAGAGGAGAACACGCCCTCGGACGGAGGTGGCCGAGCTA 383
|||||

RESULT 7
AA138848
LOCUS
DEFINITION
mq82h05.r1 Stratagene mouse melanoma (#937312) Mus musculus cDNA
clone IMAGE:585273 5' similar to gb:M95586 TRANSCRIPTION FACTOR
E2-ALPHA (HUMAN) , mRNA sequence.
ACCESSION
AA138848
VERSION
AA138848.1 GI:1701049
KEYWORDS
EST.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mammalia; Euthera; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 446)
Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M., Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R.
TITLE
The WashU-HMI Mouse EST Project
JOURNAL
Unpublished (1996)
COMMENT
Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:359921
Seq primer: -28ml3 rev1 ET from Amersham
High quality sequence stop: 443.
Location/Qualifiers
1. .446
/organism="Mus musculus"
/mol_type="mRNA"

/db_xref="taxon:10090"
/clone="IMAGE:585273"
/tissue type="melanoma"
/dev stage="M2 cells"
/lab host="SOLR (kanamycin resistant)"
/clone lib="Stratagene mouse melanoma (#937312)"
/note="Organ: skin; Vector: pBluescript SK-; Site 1: EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer: Oligo dT. From M2 cells, a highly metastatic derivative of the K-1735 (mouse) melanoma. Average insert size: 1.0 kb; Uni-ZAP XR Vector; ~5' adaptor sequence: 5' GAATTCGGCACCAG 3' ~3' adaptor sequence: 5' CTCAGTTTTTTTTTTTTTTT 3' "

Alignment Scores:
Pred. No.: 1.4e-09 Length: 446
Score: 19.00 Matches: 19
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 65.52% Indels: 0
Gaps: 1
DB:

US-10-624-218-1 (1-29) x AA138848 (1-446)

Qy 1 ArgAlaAlaPheLeuGluLysGluAenThrAlaLeuArgThrGluValAlaGluLeu 19
|||||

Db 88 CGGCAGCCTTCCTGGAGAGGAGAACACGCCCTCGGACGGAGGTGGCCGAGCTT 144
|||||

RESULT 8
BG986696
LOCUS
DEFINITION
PM1-HT1170-030101-001-b09 HT1170 Homo sapiens cDNA, mRNA sequence.
ACCESSION
BG986696
VERSION
BG986696.1 GI:14390766
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 464)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H., Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.
TITLE
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE
20202663
PUBMED
10737800
COMMENT
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=PM1&t2=PM1-HT1170-030101-001-b09&t3=2001-01-03&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 6
High quality sequence stop: 396.
Location/Qualifiers
1. .464
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev stage="Adult"
/clone_lib="HT1170"

FEATURES
source

FEATURES
source

/note="Organ: head neck; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN

Alignment Scores:
Pred. No.: 1.45e-09 Length: 464
Score: 19.00 Matches: 19
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 65.52% Indels: 0
DB: 4 Gaps: 0

US-10-624-218-1 (1-29) x BG986696 (1-464)

Qy 1 ArgAlaAlaPheLeuGluGlyGluAsnThrAlaLeuArgThrGluValAlaGluLeu 19
Db 169 CGGCAGCCTTCTCGAGAGGAGAGACACAGCCTCGCGAGGTCGCCGAGCTA 225

RESULT 9

BU262464 484 bp mRNA linear EST 26-NOV-2002
LOCUS BU262464.1 GI:25530860
DEFINITION BU262464.1 Gallus gallus cDNA clone CHEST422b16 5', mRNA sequence.

ACCESSION

VERSION BU262464.1

KEYWORDS EST.

SOURCE Gallus gallus (chicken)

ORGANISM Gallus gallus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.

REFERENCE

AUTHORS Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E., Fong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.
TITLE A Comprehensive Collection of Chicken cDNAs
JOURNAL Curr. Biol. 12 (22), 1965-1969 (2002)

MEDLINE

PUBMED 22335534

COMMENT 12445392

Contact: Simon Hubbard
Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology (UMIST)

PO Box 88, Manchester, M60 1QD, UK

Tel: 01612008930

Fax: 01612360409

Email: Simon.Hubbard@umist.ac.uk.

FEATURES

source

1..484
Location/Qualifiers
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="White Leghorn, Hisex"
/db_xref="taxon:9031"
/clone="CHEST422b16"
/dev_stage="22"
/lab_host="DH10B"
/clone_lib="CSQCHN51"

/note="Organ: limbs; Vector: pBluescript II KS(+); Site_1: EcoRI; Site_2: NotI; This normalized library was constructed from 1 million independent clones. cDNA synthesis was initiated using an oligo(dT) primer, using methylated C in the first strand synthesis reaction. Following this first strand reaction, double-stranded cDNA was blunt-ended, ligated to NotI adapters, digested with EcoRI, size-selected, and cloned into the NotI and EcoRI compatible sites of a custom modified MCS of the pBluescript (KS+) vector. The library was normalized in 2 rounds using conditions adapted from Soares et al., PNAS (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6

(1996): 791, except that a significantly longer reannealing hybridization was used."

ORIGIN

Alignment Scores:
Pred. No.: 1.51e-09 Length: 484
Score: 19.00 Matches: 19
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 65.52% Indels: 0
DB: 5 Gaps: 0

US-10-624-218-1 (1-29) x BU262464 (1-484)

Qy 1 ArgAlaAlaPheLeuGluGlyGluAsnThrAlaLeuArgThrGluValAlaGluLeu 19
Db 261 CGGCAGCCTTCTCGAGAGGAGAGATACGGCCTCGAGAGCGAGGTTGCAGAGCTG 317

RESULT 10

BF660216

LOCUS BF660216.1

DEFINITION BF660216.1 GI:11925350

ACCESSION BF660216

VERSION BF660216.1

KEYWORDS EST.

SOURCE Mus musculus

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

REFERENCE 1 (bases 1 to 528)

AUTHORS NCI-CCGAP

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

JOURNAL Unpublished (1997)

COMMENT Other ESTs: maa29a04.x1

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Jeffrey E. Green, M.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CCGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

MGI:1454374

Seq primer: -40RP from Gibco

High quality sequence stop: 435.

Location/Qualifiers

1..528

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

/clone="IMAGE:3812262"

/sex="female"

/dev_stage="10 weeks"

/lab_host="DH10B (T1 phage-resistant)"

/clone_lib="NCI CGAP L110"

/note="Organ: liver; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.

Average insert size 1.6 kb. Library constructed by Life Technologies."

ORIGIN

Alignment Scores:
Pred. No.: 1.64e-09 Length: 528
Score: 19.00 Matches: 19
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 65.52% Indels: 0
DB: 2 Gaps: 0

US-10-624-218-1 (1-29) x BR660216 (1-528)	
Qy	1 ArgAlaAlaPheLeuGluLyysGluAsnThrAlaLeuArgThrGluValAlaGluLeu 19
Db	436 CGGCGAGCCTTCTCGGAGAGGAGAACACAGCCCTGCGGACGGAGGTTCGCGAGCTT 492
RESULT 11	
CRS36622	540 bp mRNA linear EST 06-JUL-2004
LOCUS	DKFZp459A102.r1.459 (synonym: pcor1) Pongo pygmaeus cDNA clone
DEFINITION	DKFZp459A102.5', mRNA sequence.
ACCESSION	CRS36622
VERSION	CRS36622.1 GI:49683760
KEYWORDS	EST.
SOURCE	Pongo pygmaeus (orangutan)
ORGANISM	Pongo pygmaeus
REFERENCE	
AUTHORS	
1 (bases 1 to 540)	
Wambutt, R., Heubner, D., Mewes, H.W., Weil, B., Amid, C., Osanger, A., Fodor, G., Han, M. and Wiemann, S.	
Pongo pygmaeus mRNA (Wambutt, R., Heubner, D., Mewes, H.W., et al.)	
Unpublished (2004)	
Contact: MIPS	
MIPS	
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany	
This is the 5' sequence of the clone insert. Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKF2); Email s.wiemann@dkfz-heidelberg.de; sequenced by Agowa (Berlin/Germany) within the cDNA sequencing consortium of the German Genome Project. This clone (DKF2p459A102) is available at the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further information about the clone and the sequencing project is available at http://mips.gsf.de/projects/cdna/	
FEATURES	
source	
1..540	
/organism="Pongo pygmaeus"	
/mol_type="mRNA"	
/db_xref="taxon:9600"	
/clone="DKFZp459A102"	
/tissue_type="cortex"	
/dev_stage="adult"	
/lab_host="DH10B"	
/clone_lib="459 (synonym: pcor1)"	
/note="Vector: pSport1_Sfi; Site_1: SfiI; Site_2: SfiIb"	
ORIGIN	
Alignment Scores:	
Pred. No.:	1..678-09 Length: 540
Score:	19.00 Matches: 19
Percent Similarity:	100.00% Conservative: 0
Best Local Similarity:	100.00% Mismatches: 0
Query Match:	65.52% Indels: 0
DB:	7 Gaps: 0
US-10-624-218-1 (1-29) x CRS36622 (1-540)	
Qy	1 ArgAlaAlaPheLeuGluLyysGluAsnThrAlaLeuArgThrGluValAlaGluLeu 19
Db	337 CGGCGGCGCTTCTCGGAGAGGAGAACACAGCCCTGCGGACGGAGGTTCGCGAGCTA 393
RESULT 12	
AW836791/c	555 bp mRNA linear EST 18-MAY-2000
LOCUS	QV1-LT0036-150200-070-f06 LT0036 Homo sapiens cDNA, mRNA sequence.
DEFINITION	AW836791
ACCESSION	AW836791
VERSION	AW836791.1 GI:7930765
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	
AUTHORS	
1 (bases 1 to 557)	
Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E.,	
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
1 (bases 1 to 555)	
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,	
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,	
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,	
Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V.,	
O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and	
Simpson, A.J.	
Shotgun sequencing of the human transcriptome with ORF expressed	
sequence tags	
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)	
20202663	
PUBMED	
10737800	
COMMENT	
Contact: Simpson A.J.G.	
Laboratory of Cancer Genetics	
Ludwig Institute for Cancer Research	
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,	
Brazil	
Tel: +55-11-2704922	
Fax: +55-11-2707001	
Email: asimpson@ludwig.org.br	
This sequence was derived from the FAPESP/LICR Human Cancer Genome	
Project. This entry can be seen in the following URL	
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t2=QV1-LT0036-150	
200-070-f06&t3=2000-02-15&t4=1)	
Seq primer: puc 18 forward	
High quality sequence stop: 555.	
FEATURES	
source	
1..555	
/organism="Homo sapiens"	
/mol_type="mRNA"	
/db_xref="taxon:9606"	
/dev_stage="Adult"	
/clone_lib="LT0036"	
/note="Organ: leiomyos; Vector: puc18; Site 1: SmaI;	
Site 2: SmaI; A mini-library was made by cloning products	
derived from ORESTES PCR (U.S. Letters Patent application	
No. 196,716 - Ludwig Institute for Cancer Research)	
profiles into the pUC 18 vector. Reverse transcription of	
tissue mRNA and cDNA amplification were performed under	
low stringency conditions."	
ORIGIN	
Alignment Scores:	
Pred. No.:	1..72e-09 Length: 555
Score:	19.00 Matches: 19
Percent Similarity:	100.00% Conservative: 0
Best Local Similarity:	100.00% Mismatches: 0
Query Match:	65.52% Indels: 0
DB:	2 Gaps: 0
US-10-624-218-1 (1-29) x AW836791 (1-555)	
Qy	1 ArgAlaAlaPheLeuGluLyysGluAsnThrAlaLeuArgThrGluValAlaGluLeu 19
Db	368 CGGCGAGCCTTCTCGGAGAGGAGAACACAGCCCTGCGGACGGAGGTTCGCGAGCTA 312
RESULT 13	
BU311019	557 bp mRNA linear EST 28-NOV-2002
LOCUS	603540871.F1 CSEQCHN61 Gallus gallus cDNA clone CHES7509n17.5', mRNA
DEFINITION	sequence.
ACCESSION	BU311019
VERSION	BU311019.1 GI:25819020
KEYWORDS	EST.
SOURCE	Gallus gallus (chicken)
ORGANISM	Gallus gallus
REFERENCE	
AUTHORS	
1 (bases 1 to 557)	
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;	
Phasianinae; Gallus.	
1 (bases 1 to 557)	
Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E.,	

Fong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.

A Comprehensive Collection of Chicken cDNAs
Curr. Biol. 12 (22), 1965-1969 (2002)

22335534

TITLE

JOURNAL

MEDLINE

PUBMED

COMMENT

Contact: Simon Hubbard
Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology
(UMIST)

PO Box 88, Manchester, M60 1QD, UK

Tel: 01612008930

Fax: 01612360409

Email: Simon.Hubbard@umist.ac.uk.

FEATURES

source

Location/Qualifiers
1..557
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="Compton Line 151"
/db_xref="taxon:9031"
/clone="CHEST509n17"
/sex="Female"
/dev_stage="adult"
/lab_host="DH108"
/clone_lib="CSQCHN61"
/note="Organ: heart; Vector: pBluescript II KS(+); Site_1:
EcoRI; Site_2: NotI; This normalized library was
constructed from 1 million independent clones. cDNA
synthesis was initiated using an oligo(dT) primer, using
methylated C in the first strand synthesis reaction.
Following this first strand reaction, double-stranded cDNA
was blunted, ligated to NotI adapters, digested with
EcoRI, size-selected, and cloned into the NotI and EcoRI
compatible sites of a custom modified MCS of the
pBluescript (KS+) vector. The library was normalized in 2
rounds using conditions adapted from Soares et al., PNAS
(1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
(1996): 791, except that a significantly longer
reannealing hybridization was used."

ORIGIN

Alignment Scores:
Pred. No.: 1.72e-09 Length: 557
Score: 19.00 Matches: 19
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 65.52% Indels: 0
DB: 5 Gaps: 0

US-10-624-218-1 (1-29) x BU311019 (1-557)

QY 1 ArgAlaAlaPheLeuGluLysGluAenThrAlaLeuArgThrGluValAlaGluLeu 19
|||||
Db 117 CGGGCAGCCTTCTCTGAGAAAGAGATACGGCCCTGAGCGGAGGTTCAGAGCTG 173

RESULT 14

AL120537

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1 (bases 1 to 578)

EST (Blum, et al.)

Unpublished (1999)

Contact: MIPS

MIPS

Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany

AL120537 578 bp mRNA linear EST 04-SEP-2003

DKFP2761A019_r1 761 (synonym: hamy2) Homo sapiens cDNA clone

DKFP2761A019_5', mRNA sequence.

AL120537

AL120537.1 GI:5926436

EST.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 578)

EST (Blum, et al.)

Unpublished (1999)

Contact: MIPS

MIPS

Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany

This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
Sequenced by LMU (Ludwig Maximilians University,
Munich/Germany) within the cDNA sequencing consortium of the German
Genome Project.

No sl sequence available.

This clone (DKFP2761A019) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES

source

Location/Qualifiers
1..578
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFP2761A019"
/tissue_type="amygdala"
/dev_stage="adult"
/lab_host="DH108"
/clone_lib="761 (synonym: hamy2)"
/note="Vector: pSport1; Site_1: NotI; Site_2: SalI"

ORIGIN

Alignment Scores:
Pred. No.: 1.78e-09 Length: 578
Score: 19.00 Matches: 19
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 65.52% Indels: 0
DB: 1 Gaps: 0

US-10-624-218-1 (1-29) x AL120537 (1-578)

QY 1 ArgAlaAlaPheLeuGluLysGluAenThrAlaLeuArgThrGluValAlaGluLeu 19
|||||
Db 370 CGGGCAGCCTTCTCTGAGAAAGAGAACACAGCCCTGCGGACGAGGTTCAGAGCTA 426

RESULT 15

AW500333

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1 (bases 1 to 599)

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Eco RI site shown at the beginning of the sequence.

Tissue Procurement: Louis M. Staudt, M.D., Ph.D.

cDNA Library Preparation: M.B. Soares Lab

cDNA Library Arrayed by: M.B. Soares Lab

DNA Sequencing by: M.B. Soares Lab

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

www.bio.llnl.gov/bbrp/image/image.html

Seq primer: M13 Forward.

Location/Qualifiers

1..599

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:3077127"

/tissue_type="lymph"

/cell_type="germinal center B cells"

/cell_line="MGC85"

/lab_host="DH10B (LTI)"
/clone_lib="NIH_MGC 50"
/note="vector: pT73-Pac; Site_1: NotI; Site_2: Eco RI;
Constructed from size fractionated cytoplasmic mRNA
(3.5-4.4kb). Directionally cloned. Cells provided by
Louis M. Staudt, Ph.D. Library preparation by Maria de
Fatima Bonaldo, Ph.D. and M. Bento Soares, Ph.D."

ORIGIN

Alignment Scores:
Pred. No.: 1.84e-09 Length: 599
Score: 19.00 Matches: 19
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 65.52% Indels: 0
DB: 2 Gaps: 0

US-10-624-218-1 (1-29) x AW500333 (1-599)

Qy 1 ArgAlaAlaPheLeuGluLysGluAsnThrAlaLeuArgThrGluValAlaGluLeu 19
Db 166 CGGGCAGCCTTCCTGGAGAGGAGACACAGCCCTGCGGACGGAGGTGGCCGAGCTA 222

Search completed: July 27, 2005, 16:41:08
Job time : 2025.31 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 26, 2005, 11:54:19 ; Search time 73.3055 Seconds
(without alignments)
153.004 Million cell updates/sec

Title: US-10-624-218-2
Perfect score: 29
Sequence: 1 RAAFLEKENTALTRVABLRKRVGRCRNI 29

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 2105692 seqs, 386760381 residues

Word size : 0
Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

- Database : A_Geneseq_16Dec04:*
- 1: Geneseqp1980s:*
 - 2: Geneseqp1990s:*
 - 3: Geneseqp2000s:*
 - 4: Geneseqp2001s:*
 - 5: Geneseqp2002s:*
 - 6: Geneseqp2003as:*
 - 7: Geneseqp2003bs:*
 - 8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	29	100.0	29	ADI34277	ADI34277 Zipper pe
2	14	48.3	29	ADI34276	ADI34276 Zipper pe
3	14	48.3	261	ADD47845	ADD47845 Rat Prote
4	14	48.3	303	ADD47847	ADD47847 Human Pro
5	14	48.3	303	ADJ68701	ADJ68701 Human hea
6	14	48.3	303	ADI82509	ADI82509 Human mod
7	14	48.3	303	ABM82367	ABM82367 Tumour-as
8	9	31.0	325	ADSE6596	ADSE6596 Human Pro
9	9	31.0	325	ADD46912	ADD46912 Rat Prote
10	9	31.0	325	ADSE6594	ADSE6594 Rat Prote
11	9	31.0	325	ADD46916	ADD46916 Rat Prote
12	9	31.0	325	ADD46914	ADD46914 Human Pro
13	9	31.0	325	ADD46918	ADD46918 Human Pro
14	9	31.0	325	ADSE6590	ADSE6590 Rat Prote
15	9	31.0	325	ADSE6592	ADSE6592 Human Pro
16	9	31.0	325	ADI82507	ADI82507 Human mod
17	9	31.0	325	ADDI19734	ADDI19734 Human PRO
18	9	31.0	325	ADP54859	ADP54859 Human PRO
19	8	27.6	498	AAB79524	AAB79524 Corynebact
20	8	27.6	539	ADL90009	ADL90009 Gluconoba
21	8	27.6	675	AAG92163	AAG92163 C glutami
22	8	27.6	675	ADDI13431	ADDI13431 C. glutam
23	7	24.1	32	AAB61569	AAB61569 Peptide W
24	7	24.1	32	AAB61565	AAB61565 Peptide W
25	7	24.1	43	ADI34278	ADI34278 Zipper pe

26	7	24.1	249	6	ABU24158	Abu24158 Protein e
27	7	24.1	375	7	ADG48189	Adg48189 Mycobacte
28	7	24.1	401	6	ABP70989	Abp70989 Epoxide h
29	7	24.1	765	8	ADN46635	Adn46635 Thermococ
30	7	24.1	767	5	ABP35613	Abp35613 Fungal ZB
31	7	24.1	777	4	AAB96565	Aab96565 Putative
32	7	24.1	790	8	ADS43080	Ads43080 Bacterial
33	6	20.7	35	5	ABG70818	Abg70818 Mouse myo
34	6	20.7	35	6	ABG74186	Abg74186 Mouse myo
35	6	20.7	43	8	ADI34279	Adi34279 Zipper pe
36	6	20.7	54	5	ADK35905	Adk35905 Novel hum
37	6	20.7	72	4	AU56794	Au56794 Propionib
38	6	20.7	72	6	ABM53313	Abm53313 Propionib
39	6	20.7	80	4	AU62784	Au62784 Propionib
40	6	20.7	80	6	ABM59303	Abm59303 Propionib
41	6	20.7	81	6	ADB06462	Adb06462 Allolococ
42	6	20.7	85	6	ADB06464	Adb06464 Allolococ
43	6	20.7	86	7	ADC94503	Adc94503 E. faeciu
44	6	20.7	95	3	ABB06292	Abb06292 Streptomy
45	6	20.7	98	3	AAG10672	Aag10672 Arabidops

ALIGNMENTS

RESULT 1
ADI34277
ID ADI34277 standard; protein; 29 AA.
XX
AC ADI34277;
XX
DT 15-APR-2004 (first entry)
XX
DE Zipper peptide #2 for cross linking adenoviral ligands.
XX
KW adenovirus vector; Cytostatic; HER2/neu; CD40; tumor necrosis factor;
TNF.
XX
OS Unidentified.
XX
PN WO2004009133-A1.
XX
PD 29-JAN-2004.
XX
PF 22-JUL-2003; 2003WO-US022852.
XX
PR 22-JUL-2002; 2002US-0397951P.
XX
PS (VECT-) VECTORLOGICS INC.
XX
PI Korokhov N, Mikheeva G;
XX
DR WPI; 2004-132871/13.
XX
PT Novel recombinant adenovirus having fiber protein modified by insertion
of first zipper peptide that can crosslink to second zipper peptide-
targeting ligand fusion protein, and binding between zipper peptides
targets vector to cell.
XX
PS Claim 2; SEQ ID NO 2; 54pp; English.

The present invention relates to a targeted recombinant adenovirus vector. The invention is useful for expressing a heterologous protein chosen from a tumor associated antigen, HER2/neu and carcinoembryonic antigen, in a target cell e.g., CD40 + cell such as dendritic cells, is useful in a method of gene transfer to CD40 + cells, where the targeted adenovirus vector mediates transfer of the gene encoding heterologous protein to the cell such as the dendritic cell. The vector is useful in gene therapy techniques for treatment of tumors. Multivalent interaction or trimeric CD40L with CD40 receptors causes CD40 ligation, which then results in enhanced survival (IL)-1, IL-6, IL-8, IL-10, IL-12, tumor necrosis factor (TNF)-4agr;,, macrophage inflammatory protein (MIP)-1a and enzymes

CC such as matrix metalloproteinase. CD40-CD40L interaction also enhances
CC monocyte tumoricidal activity. The present sequence represents a zipper
CC peptide to be used for cross-linking targeting ligands with adenoviral
CC virions.
XX
XX
SQ Sequence 29 AA;

Query Match 100.0%; Score 29; DB 8; Length 29;
Best Local Similarity 100.0%; Pred. No. 3.4e-22;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RAAFLKENTALTTRVAELRRKRGCRNI 29
Db 1 RAAFLKENTALTTRVAELRRKRGCRNI 29

RESULT 2
ADI34276
ID ADI34276 standard; protein; 29 AA.
XX
AC ADI34276;
XX
DT 15-APR-2004 (first entry)
XX
DE Zipper peptide #1 for cross linking adenoviral ligands.
XX
KW adenovirus vector; Cytostatic; HER2/neu; CD40; tumor necrosis factor;
KW TNF.
XX
XX Unidentified.
OS
XX WO2004009133-A1.
PN
XX 29-JAN-2004.
PD
XX
XX 22-JUL-2003; 2003WO-US022852.
PF
XX 22-JUL-2002; 2002US-0397951P.
PR
XX (VECT-) VECTORLOGICS INC.
PA
XX
XX Korokhov N, Mikheeva G;
PI
XX
XX WPI; 2004-132871/13.
DR
XX
XX Novel recombinant adenovirus having fiber protein modified by insertion
PT of first zipper peptide that can crosslink to second zipper peptide-
PT targeting ligand fusion protein, and binding between zipper peptides
PT targets vector to cell.
XX
XX Claim 2; SEQ ID NO 1; 54pp; English.
PS
XX
XX The present invention relates to a targeted recombinant adenovirus
CC vector. The invention is useful for expressing a heterologous protein
CC chosen from a tumor associated antigen, HER2/neu and carcinoembryonic
CC antigen, in a target cell e.g., CD40 + cell such as dendritic cells, is
CC useful in a method of gene transfer to CD40 + cells, where the targeted
CC adenovirus vector mediates transfer of the gene encoding heterologous
CC protein to the cell such as the dendritic cell. The vector is useful in
CC gene therapy techniques for treatment of tumors. Multivalent interaction
CC or trimeric CD40L with CD40 receptors causes CD40 ligation, which then
CC results in enhanced survival of these cells and secretion of cytokines
CC such as interleukin (IL)-1, IL-6, IL-8, IL-10, IL-12, tumor necrosis
CC factor (TNF)-46agrt; macrophage inflammatory protein (MIP)-1a and enzymes
CC such as matrix metalloproteinase. CD40-CD40L interaction also enhances
CC monocyte tumoricidal activity. The present sequence represents a zipper
CC peptide to be used for cross-linking targeting ligands with adenoviral
CC virions.
XX
XX Sequence 29 AA;
SQ

Query Match 48.3%; Score 14; DB 8; Length 29;
Best Local Similarity 100.0%; Pred. No. 6e-07;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RAAFLKENTALT 14
Db 1 RAAFLKENTALT 14

RESULT 3
ADD47845
ID ADD47845 standard; protein; 261 AA.
XX
AC ADD47845;
XX
DT 02-DEC-2004 (revised)
DT 29-JAN-2004 (first entry)
XX
DE Rat Protein AAB20032, SEQ ID NO 13541.
XX
XX Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
XX
XX Rattus norvegicus.
OS Unidentified.
XX
XX WO2003016475-A2.
PN
XX 27-FEB-2003.
PD
XX 14-AUG-2002; 2002WO-US025765.
PF
XX 14-AUG-2001; 2001US-0312147P.
PR
XX 01-NOV-2001; 2001US-0346382P.
PR
XX 26-NOV-2001; 2001US-033347P.
XX
XX (GEO) GEN HOSPITAL CORP.
PA (FARB) BAYER AG.
PA
XX Woolf C, D'urso D, Befort K, Costigan M;
PI
XX WPI; 2003-268312/26.
DR GENBANK; AAB20032.
XX
XX New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
PT
XX Example 1; Page; 1017pp; English.
PS
XX
XX The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative of allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. Gene
CC therapy). The sequence presented is a rat protein (described in Table 3
CC of the specification) which is differentially expressed during pain.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at

CC ftp.wipo.int/pub/published_pct_sequences.
 XX Sequence 261 AA;
 SQ

Query Match 48.3%; Score 14; DB 7; Length 261;
 Best Local Similarity 100.0%; Pred. No. 4.2e-06;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RAAFLKENTALT 14
 |||||
 DB 222 RAAFLKENTALT 235

RESULT 4
 ID ADD47847 standard; protein; 303 AA.
 XX ADD47847;
 AC

02-DEC-2004 (revised)
 29-JAN-2004 (first entry)
 XX

Human Protein NP_003207, SEQ ID NO 13543.
 DE

Human; pain; neuronal tissue; gene therapy;
 spinal segmental nerve injury; chronic constriction injury; CCI;
 spared nerve injury; SNI; Chung.
 KW
 KW
 KW
 XX

Homo sapiens.
 OS
 Unidentified.
 OS

WO2003016475-A2.
 PN

27-FEB-2003.
 XX

14-AUG-2002; 2002WO-US025765.
 PF

14-AUG-2001; 2001US-0312147P.
 PR

01-NOV-2001; 2001US-0346382P.
 PR

26-NOV-2001; 2001US-0333347P.
 PR

(GEO) GEN HOSPITAL CORP.
 PA
 (FARB) BAYER AG.
 PA

Woolf C, D'urso D, Befort K, Costigan M;
 PI

WPI; 2003-268312/26.
 DR

GENBANK; NP_003207.
 DR

New composition comprising two or more isolated polypeptides, useful for
 preparing a medicament for treating pain in an animal.
 PT

Example 1; Page; 1017pp; English.
 PS

The invention discloses a composition comprising two or more isolated rat
 or human polynucleotides or a polynucleotide which represents a fragment,
 CC
 CC derivative or allelic variation of the nucleic acid sequence. Also
 CC
 CC claimed are a vector comprising the novel polynucleotide, a host cell
 CC comprising the vector, a method for identifying a nucleotide sequence
 CC which is differentially regulated in an animal subjected to pain and a
 CC kit to perform the method, an array, a method for identifying an agent
 CC that increases or decreases the expression of the polynucleotide sequence
 CC that is differentially expressed in neuronal tissue of a first animal
 CC subjected to pain, a method for identifying a compound which regulates
 CC the expression of a polynucleotide sequence which is differentially
 CC expressed in an animal subjected to pain, a method for identifying a
 CC compound that regulates the activity of one or more of the
 CC polynucleotides, a method for producing a pharmaceutical composition, a
 CC method for identifying a compound or small molecule that regulates the
 CC activity in an animal of one or more of the polypeptides given in the
 CC specification, a method for identifying a compound useful in treating
 CC pain and a pharmaceutical composition comprising the one or more
 CC polypeptides or their antibodies. The polynucleotide or the compound that

modulates its activity is useful for preparing a medicament for treating
 pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
 injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
 therapy). The sequence presented is a human protein (described in Table 3
 of the specification) which is differentially expressed during pain.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic form directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX

Sequence 303 AA;
 SQ

Query Match 48.3%; Score 14; DB 7; Length 303;
 Best Local Similarity 100.0%; Pred. No. 4.7e-06;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RAAFLKENTALT 14
 |||||
 DB 264 RAAFLKENTALT 277

RESULT 5
 ADJ68701

ID ADJ68701 standard; protein; 303 AA.
 XX

ADJ68701;
 AC

06-MAY-2004 (first entry)
 DT

Human heat mitochondrial protein as a therapeutic target SeqID507.
 DE

mitochondrial; human; screening assay; diabetes mellitus;
 Huntington's disease; osteoarthritis;
 Leber's hereditary optic neuropathy; LHON;
 KW mitochondrial encephalopathy lactic acidosis and stroke; MELAS;
 KW myoclonic epilepsy; tagged red fibre syndrome; MERRF; cancer;
 KW neuroprotective; nontropic; antidiabetic; anticonvulsant; antiarthritic;
 KW osteopathic; ophthalmological; cytostatic.
 KW

Homo sapiens.
 OS

WO2003087768-A2.
 PN

23-OCT-2003.
 PD

04-APR-2003; 2003WO-US010870.
 PF

12-APR-2002; 2002US-0372843P.
 PR

17-JUN-2002; 2002US-0389987P.
 PR

20-SEP-2002; 2002US-0412418P.
 PR

(MITO-) MITOKOR.
 PA
 (BUCK-) BUCK INST AGE RES.
 PA

Ghosh SS, Fahy ED, Zhang B, Gibson BW, Taylor SW, Glenn GM;
 PI Warnock DE;
 PI

WPI; 2003-845369/78.
 DR

Identifying a mitochondrial target for drug screening assays and for
 treating diseases associated with altered mitochondrial function,
 PT comprises detecting a modified polypeptide in a sample and correlating
 PT with the disease.
 PT

Claim 1; SEQ ID NO 507; 180pp; English.
 PS

This invention relates to novel mitochondrial targets that can be used
 for therapeutic intervention in treating a disease associated with
 altered mitochondrial function. Specifically, it refers to a method for
 identifying proteins of the human heart mitochondrial proteome that are
 CC useful for drug screening assays, as well as therapeutic targets. The
 CC present invention describes a method for identifying such proteins that
 CC can be used in the treatment of various diseases associated with altered
 CC mitochondrial function including diabetes mellitus, Huntington's disease,

CC osteoarthritis, Leber's hereditary optic neuropathy (LHON), mitochondrial
 CC encephalopathy lactic acidosis and stroke (MELAS), myoclonic epilepsy
 CC ragged red fibre syndrome (MERRF) or cancer. Accordingly, these
 CC compositions have neuroprotective, neurotropic, antidiabetic,
 CC anticonvulsant, antiarthritic, osteopathic, ophthalmological and
 CC cytoskeletal activities. This polypeptide sequence is a human heart
 CC mitochondrial protein of the invention.

XX Sequence 303 AA;

Query Match 48.3%; Score 14; DB 7; Length 303;
 Best Local Similarity 100.0%; Pred. No. 4.7e-06;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RAAFLKENTALT 14
 |||||
 Db 264 RAAFLKENTALT 277

RESULT 6
 ADI82509
 ID ADI82509 standard; protein; 303 AA.

XX ADI82509;

DT 22-APR-2004 (first entry)

DE Human modifier of p21 (MP21) protein sequence SeqID75.

XX p21 pathway modulating agent; assay system; MP21; cytostatic;

KW MP21 protein activity modulator; cancer; genetically modified animal;
 KW human.

XX Homo sapiens.

XX WO2004005486-A2.

XX 15-JAN-2004.

XX 09-JUL-2003; 2003WO-US021510.

XX 10-JUL-2002; 2002US-0394795P.

PR 07-AUG-2002; 2002US-0401739P.

PR 16-SEP-2002; 2002US-0411010P.

PR 30-DEC-2002; 2002US-0437158P.

XX (EXEL-) EXELIXIS INC.

XX Francis-Lang H, Friedman L, Kidd T, Roche S, Joo DM, Lickteig K;
 PI Amundsen CD, Hai B, Zhang H, Adamkewicz JI, Hammonds GR;

XX WPI; 2004-091358/09.

DR N-PSDB; ADI82448.

XX Identifying a candidate p21 pathway modulating agent, useful for treating
 PT a disease such as cancer, comprises contacting an assay system comprising
 PT a MP21 polypeptide or nucleic acid with a test agent.

XX Example 2; SEQ ID NO 75; 392pp; English.

XX This invention relates to a novel candidate p21 pathway modulating agent
 CC by contacting an assay system comprising an MP21 (modifier of p21)
 CC polypeptide or nucleic acid with a test agent, where in the absence of
 CC the test agent the system provides a reference activity and detecting a
 CC test agent-biased activity of the assay system. The invention may be
 CC useful for the production of compounds with a cytostatic activity through
 CC modulation of MP21 protein activity. The MP21 polypeptide or nucleic acid
 CC can be used for identifying MP21 modulating agents useful as therapeutic
 CC targets for diagnosing cancer or treating disorders associated with
 CC defective or impaired p21 and/or MP21 function. MP21 modulating agents
 CC are useful in diagnosis, therapy, for example treating cancer, and
 CC pharmaceutical development. The genetically modified animals may be used
 CC for in vivo assays to test for activity of a candidate p21 modulating

CC agent, or to further assess the role of MP21 in a p21 pathway process.
 CC The present sequence is that of a human MP21 protein which is an
 CC orthologue of a Drosophila p21 modifier and which was used in the
 CC exemplification of the invention.

XX Sequence 303 AA;

Query Match 48.3%; Score 14; DB 8; Length 303;
 Best Local Similarity 100.0%; Pred. No. 4.7e-06;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RAAFLKENTALT 14
 |||||
 Db 264 RAAFLKENTALT 277

RESULT 7
 ABM82367
 ID ABM82367 standard; protein; 303 AA.

XX ABM82367;

DT 18-NOV-2004 (first entry)

XX Tumour-associated antigenic target (TAT) polypeptide PR083291, SEQ:6078.

XX Tumour-associated antigenic target; TAT; human; overexpression; cancer;

KW tumour; diagnosis; cell proliferative disorder; breast cancer;

KW colorectal cancer; lung cancer; ovarian cancer; liver cancer;

KW central nervous system cancer; bladder cancer; pancreatic cancer;

KW cervical cancer; melanoma; leukaemia; hybridisation probe;

KW chromosome identification; chromosome mapping; gene mapping;

KW gene therapy; cytostatic.

XX Homo sapiens.

XX WO2004030615-A2.

XX 15-APR-2004.

XX 29-SEP-2003; 2003WO-US028547.

PR 02-OCT-2002; 2002US-0414971P.

XX (GETH) GENENTECH INC.

XX Wu TD, Zhang Z, Zhou Y;

XX WPI; 2004-347921/32.

DR N-PSDB; ACN40954.

XX New tumor-associated antigenic target polypeptides and nucleic acids,
 PT useful in preparing a medicament for treating or detecting a
 PT proliferative disorder, e.g. breast, lung, colorectal, ovarian or
 PT prostate cancer or tumor.

XX Claim 12; SEQ ID NO 6078; 7273pp; English.

XX The invention relates to human tumour-associated antigenic target (TAT)
 CC polypeptides, and their related nucleic acids. The TAT polypeptides are
 CC overexpressed in cancer tissues compared to normal tissues, and may thus
 CC serve as effective targets for the diagnosis and treatment of cancer in
 CC mammals. The invention also relates to nucleic acid and polypeptide
 CC sequences at least 80% identical to the TAT nucleic acids and
 CC polypeptides; expression vectors and host cells comprising a TAT nucleic
 CC acid; an antibody specific for a TAT polypeptide; a peptide or organic
 CC molecule which binds to a TAT polypeptide; fusion proteins comprising a
 CC TAT polypeptide; and methods and compositions for the treatment or
 CC diagnosis of cancer in mammals. TAT polypeptides, nucleic acids,
 CC antibodies, antagonists, binding molecules and compositions are useful
 CC for diagnosing or treating a cell proliferative disorder associated with
 CC increased TAT expression, particularly cancers such as breast cancer,
 CC colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder

CC cancer, pancreatic cancer, cervical cancer, cancers of the central
CC nervous system, melanoma and leukemia. TAT nucleic acids may further be
CC used as hybridisation probes, in chromosome and gene mapping, in
CC chromosome identification and in gene therapy. The present sequence
CC represents a TAT polypeptide of the invention
XX
SQ Sequence 303 AA;

Query Match 48.3%; Score 14; DB 8; Length 303;
Best Local Similarity 100.0%; Pred. No. 4.7e-06;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RAAFLKENTALRT 14
Db 264 RAAFLKENTALRT 277
|||||

RESULT 8
ADE56596
ID ADE56596 standard; protein; 325 AA.

AC ADE56596;

XX 29-JAN-2004 (first entry)

XX Human Protein Q10586, SEQ ID NO 2450.

XX Human; pain; neuronal tissue; gene therapy;
KW spinal segmental nerve injury; chronic constriction injury; CCI;
KW spared nerve injury; SNI; Chung.

XX Homo sapiens.

XX WO2003016475-A2.

XX 27-FEB-2003.

XX 14-AUG-2002; 2002WO-US025765.

XX 14-AUG-2001; 2001US-0312147P.

PR 01-NOV-2001; 2001US-0346382P.

PR 26-NOV-2001; 2001US-0333347P.

XX (GEO) GEN HOSPITAL CORP.

PA (FARB) BAYER AG.

XX Woolf C, D'urso D, Befort K, Costigan M;

PI WPI; 2003-268312/26.

XX GENBANK; Q10586.

DR New composition comprising two or more isolated polypeptides, useful for

XX preparing a medicament for treating pain in an animal.

PS Claim 1; Page; 1017pp; English.

XX The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating

CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a human protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 325 AA;

Query Match 31.0%; Score 9; DB 7; Length 325;
Best Local Similarity 100.0%; Pred. No. 0.61;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RAAFLKEN 9
Db 286 RAAFLKEN 294
|||||

RESULT 9

ADD46912
ID ADD46912 standard; protein; 325 AA.

XX ADD46912;

XX 02-DEC-2004 (revised)

DT 29-JAN-2004 (first entry)

XX Rat Protein AAA41083, SEQ ID NO 12598.

XX Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.

XX Rattus norvegicus.

OS Unidentified.

PN WO2003016475-A2.

PD 27-FEB-2003.

XX 14-AUG-2002; 2002WO-US025765.

XX 14-AUG-2001; 2001US-0312147P.

PR 01-NOV-2001; 2001US-0346382P.

PR 26-NOV-2001; 2001US-0333347P.

XX (GEO) GEN HOSPITAL CORP.

PA (FARB) BAYER AG.

XX Woolf C, D'urso D, Befort K, Costigan M;

PI WPI; 2003-268312/26.

XX GENBANK; AAA41083.

DR New composition comprising two or more isolated polypeptides, useful for

XX preparing a medicament for treating pain in an animal.

PS Example 1; Page; 1017pp; English.

XX The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially

expressed in an animal subjected to pain, a method for identifying a compound that regulates the activity of one or more of the polynucleotides, a method for producing a pharmaceutical composition, a method for identifying a compound or small molecule that regulates the activity in an animal of one or more of the polypeptides given in the specification, a method for identifying a compound useful in treating pain and a pharmaceutical composition comprising the one or more polypeptides or their antibodies. The polynucleotide or the compound that modulates its activity is useful for preparing a medicament for treating injury (CCI) and spared nerve injury (SNI) in an animal (e.g. gene therapy). The sequence presented is a rat protein (shown in Table 3 of the specification) which is differentially expressed during pain. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic form directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

Sequence 325 AA;

Query Match 31.0%; Score 9; DB 7; Length 325;
Best Local Similarity 100.0%; Pred. No. 0.61;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RAAFLFLEKEN 9
Db 286 RAAFLFLEKEN 294

RESULT 10
ADE56594
ID ADE56594 standard; protein; 325 AA.

XX ADE56594;

XX 29-JAN-2004 (first entry)

XX Rat Protein P16443, SEQ ID NO 2448.

XX Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
XX chronic constriction injury; CCI; spared nerve injury; SNI; Chung.

XX Rattus norvegicus.

XX WO2003016475-A2.

XX 27-FEB-2003.

XX 14-AUG-2002; 2002WO-US025765.

XX 14-AUG-2001; 2001US-0312147P.

XX 01-NOV-2001; 2001US-0346382P.

XX 26-NOV-2001; 2001US-0333347P.

XX (GEHO) GEN HOSPITAL CORP.

XX (FARB) BAYER AG.

XX Woolf C, D'urso D, Befort K, Costigan M;

XX WPI; 2003-268312/26.

XX GENBANK; P16443.

XX Claim 1; Page; 1017pp; English.

XX The invention discloses a composition comprising two or more isolated rat
XX or human polynucleotides or a polynucleotide which represents a fragment,
XX derivative or allelic variation of the nucleic acid sequence. Also
XX claimed are a vector comprising the novel polynucleotide, a host cell
XX comprising the vector, a method for identifying a nucleotide sequence
XX which is differentially regulated in an animal subjected to pain and a
XX kit to perform the method, an array, a method for identifying an agent

CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a rat protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 325 AA;

Query Match 31.0%; Score 9; DB 7; Length 325;
Best Local Similarity 100.0%; Pred. No. 0.61;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RAAFLFLEKEN 9
Db 286 RAAFLFLEKEN 294

RESULT 11

ADD46916

ID ADD46916 standard; protein; 325 AA.

XX ADD46916;

XX 02-DEC-2004 (revised)

XX 29-JAN-2004 (first entry)

XX Rat Protein AAA41083, SEQ ID NO 12602.

XX Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
XX chronic constriction injury; CCI; spared nerve injury; SNI; Chung.

XX Rattus norvegicus.

XX Unidentified.

XX WO2003016475-A2.

XX 27-FEB-2003.

XX 14-AUG-2002; 2002WO-US025765.

XX 14-AUG-2001; 2001US-0312147P.

XX 01-NOV-2001; 2001US-0346382P.

XX 26-NOV-2001; 2001US-0333347P.

XX (GEHO) GEN HOSPITAL CORP.

XX (FARB) BAYER AG.

XX Woolf C, D'urso D, Befort K, Costigan M;

XX WPI; 2003-268312/26.

XX GENBANK; AAA41083.

XX New composition comprising two or more isolated polypeptides, useful for
XX preparing a medicament for treating pain in an animal.
XX Example 1; Page; 1017pp; English.
XX The invention discloses a composition comprising two or more isolated rat

CC or human polynucleotides or a polynucleotide which represents a fragment,
 CC derivative or allelic variation of the nucleic acid sequence. Also
 CC claimed are a vector comprising the novel polynucleotide, a host cell
 CC comprising the vector, a method for identifying a nucleotide sequence
 CC which is differentially regulated in an animal subjected to pain and a
 CC kit to perform the method, an array, a method for identifying an agent
 CC that increases or decreases the expression of the polynucleotide sequence
 CC that is differentially expressed in neuronal tissue of a first animal
 CC subjected to pain, a method for identifying a compound which regulates
 CC the expression of a polynucleotide sequence which is differentially
 CC expressed in an animal subjected to pain, a method for identifying a
 CC compound that regulates the activity of one or more of the
 CC polynucleotides, a method for producing a pharmaceutical composition, a
 CC method for identifying a compound or small molecule that regulates the
 CC activity in an animal of one or more of the polypeptides given in the
 CC specification, a method for identifying a compound useful in treating
 CC pain and a pharmaceutical composition comprising the one or more
 CC polypeptides or their antibodies. The polynucleotide or the compound that
 CC modulates its activity is useful for preparing a medicament for treating
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
 CC therapy). The sequence presented is a rat protein (described in Table 3
 CC of the specification) which is differentially expressed during pain.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic form directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 325 AA;

Query Match 31.0%; Score 9; DB 7; Length 325;

Best Local Similarity 100.0%; Pred. No. 0.61;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RAAFLKEN 9

DB 286 RAAFLKEN 294

RESULT 12

ADD46914
 ID ADD46914 standard; protein; 325 AA.

XX AC ADD46914;

XX DT 02-DEC-2004 (revised)

XX DT 29-JAN-2004 (first entry)

XX DE Human Protein NP_001343, SEQ ID NO 12600.

XX KW Human; pain; neuronal tissue; gene therapy;

XX KW spinal segmental nerve injury; chronic constriction injury; CCI;

XX KW spared nerve injury; SNI; Chung.

XX OS Homo sapiens.

XX OS Unidentified.

XX PN WO2003016475-A2.

XX PD 27-FEB-2003.

XX PF 14-AUG-2002; 2002WO-US025765.

XX PR 14-AUG-2001; 2001US-0312147P.

XX PR 01-NOV-2001; 2001US-0346382P.

XX PR 26-NOV-2001; 2001US-0333347P.

XX PR (GEHO) GEN HOSPITAL CORP.

XX PA (FARB) BAYER AG.

XX PI Woolf C, D'urso D, Befort K, Costigan M;

XX PR WPI; 2003-268312/26.

XX DR GENBANK; NP_001343.

XX New composition comprising two or more isolated polypeptides, useful for
 PT preparing a medicament for treating pain in an animal.

XX Example 1; Page; 1017pp; English.

XX The invention discloses a composition comprising two or more isolated rat
 CC or human polynucleotides or a polynucleotide which represents a fragment,
 CC derivative or allelic variation of the nucleic acid sequence. Also
 CC claimed are a vector comprising the novel polynucleotide, a host cell
 CC comprising the vector, a method for identifying a nucleotide sequence
 CC which is differentially regulated in an animal subjected to pain and a
 CC kit to perform the method, an array, a method for identifying an agent
 CC that increases or decreases the expression of the polynucleotide sequence
 CC that is differentially expressed in neuronal tissue of a first animal
 CC subjected to pain, a method for identifying a compound which regulates
 CC the expression of a polynucleotide sequence which is differentially
 CC expressed in an animal subjected to pain, a method for identifying a
 CC compound that regulates the activity of one or more of the
 CC polynucleotides, a method for producing a pharmaceutical composition, a
 CC method for identifying a compound or small molecule that regulates the
 CC activity in an animal of one or more of the polypeptides given in the
 CC specification, a method for identifying a compound useful in treating
 CC pain and a pharmaceutical composition comprising the one or more
 CC polypeptides or their antibodies. The polynucleotide or the compound that
 CC modulates its activity is useful for preparing a medicament for treating
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
 CC therapy). The sequence presented is a human protein (described in Table 3
 CC of the specification) which is differentially expressed during pain.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic form directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 325 AA;

Query Match 31.0%; Score 9; DB 7; Length 325;

Best Local Similarity 100.0%; Pred. No. 0.61;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RAAFLKEN 9

DB 286 RAAFLKEN 294

RESULT 13

ADD46918

ID ADD46918 standard; protein; 325 AA.

XX AC ADD46918;

XX DT 02-DEC-2004 (revised)

XX DT 29-JAN-2004 (first entry)

XX DE Human Protein NP_001343, SEQ ID NO 12604.

XX KW Human; pain; neuronal tissue; gene therapy;

XX KW spinal segmental nerve injury; chronic constriction injury; CCI;

XX KW spared nerve injury; SNI; Chung.

XX OS Homo sapiens.

XX OS Unidentified.

XX PN WO2003016475-A2.

XX PD 27-FEB-2003.

XX PF 14-AUG-2002; 2002WO-US025765.

XX PR 14-AUG-2001; 2001US-0312147P.

XX PR 01-NOV-2001; 2001US-0346382P.

XX PR 26-NOV-2001; 2001US-0333347P.

XX

```
PA (GEHO ) GEN HOSPITAL CORP.
PA (FARB ) BAYER AG.
XX
PI Woolf C, D'urso D, Befort K, Costigan M;
XX
DR WPI; 2003-268312/26.
DR GENBANK; NP_001343.
XX
PT New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
XX
PS Example 1; Page; 1017pp; English.
XX
CC The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a human protein (described in Table 3
CC of the specification) which is differentially expressed during pain.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 325 AA;

Query Match 31.0%; Score 9; DB 7; Length 325;
Best Local Similarity 100.0%; Pred. No. 0.61;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RAAFLXEN 9
Db 286 RAAFLXEN 294
|||||

RESULT 14
ADE56590
ID ADE56590 standard; protein; 325 AA.
XX
AC ADE56590;
XX
DT 29-JAN-2004 (first entry)
XX
DE Rat Protein P16443, SEQ ID NO 2444.
XX
KW Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
XX
OS Rattus norvegicus.
XX
PN WO2003016475-A2.
XX
PD 27-FEB-2003.
XX
PF 14-AUG-2002; 2002WO-US025765.
XX

14-AUG-2001; 2001US-0312147P.
01-NOV-2001; 2001US-0346382P.
26-NOV-2001; 2001US-0333347P.
(GEHO ) GEN HOSPITAL CORP.
(FARB ) BAYER AG.
Woolf C, D'urso D, Befort K, Costigan M;
WPI; 2003-268312/26.
GENBANK; P16443.
New composition comprising two or more isolated polypeptides, useful for
preparing a medicament for treating pain in an animal.
Claim 1; Page; 1017pp; English.
The invention discloses a composition comprising two or more isolated rat
or human polynucleotides or a polynucleotide which represents a fragment,
derivative or allelic variation of the nucleic acid sequence. Also
claimed are a vector comprising the novel polynucleotide, a host cell
comprising the vector, a method for identifying a nucleotide sequence
which is differentially regulated in an animal subjected to pain and a
kit to perform the method, an array, a method for identifying an agent
that increases or decreases the expression of the polynucleotide sequence
that is differentially expressed in neuronal tissue of a first animal
subjected to pain, a method for identifying a compound which regulates
the expression of a polynucleotide sequence which is differentially
expressed in an animal subjected to pain, a method for identifying a
compound that regulates the activity of one or more of the
polynucleotides, a method for producing a pharmaceutical composition, a
method for identifying a compound or small molecule that regulates the
activity in an animal of one or more of the polypeptides given in the
specification, a method for identifying a compound useful in treating
pain and a pharmaceutical composition comprising the one or more
polypeptides or their antibodies. The polynucleotide or the compound that
modulates its activity is useful for preparing a medicament for treating
pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
therapy). The sequence presented is a human protein (described in Table 3
of the specification) which is differentially expressed during pain.
Note: The sequence data for this patent did not form part of the printed
specification, but was obtained in electronic form directly from WIPO at
ftp.wipo.int/pub/published_pct_sequences.
Sequence 325 AA;

Query Match 31.0%; Score 9; DB 7; Length 325;
Best Local Similarity 100.0%; Pred. No. 0.61;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RAAFLXEN 9
Db 286 RAAFLXEN 294
|||||

RESULT 15
ADE56592
ID ADE56592 standard; protein; 325 AA.
XX
AC ADE56592;
XX
DT 29-JAN-2004 (first entry)
XX
DE Human Protein Q10586, SEQ ID NO 2446.
XX
KW Human; pain; neuronal tissue; gene therapy;
KW spinal segmental nerve injury; CCI; chronic constriction injury; CCI;
KW spared nerve injury; SNI; Chung.
XX
OS Homo sapiens.
XX
PN WO2003016475-A2.
```

XX	27-FEB-2003.
PD	
XX	14-AUG-2002; 2002WO-US025765.
XX	
PP	
XX	14-AUG-2001; 2001US-0312147P.
PR	01-NOV-2001; 2001US-0346382P.
PR	26-NOV-2001; 2001US-0333347P.
XX	
XX	(GEHO) GEN HOSPITAL CORP.
PA	(FARB) BAYER AG.
XX	
XX	Woolf C, D'urso D, Befort K, Costigan M;
PI	
XX	WPI; 2003-269312/26.
DR	GENBANK; Q10586.
DR	
XX	
XX	New composition comprising two or more isolated polypeptides, useful for
PT	preparing a medicament for treating pain in an animal.
PT	
XX	Claim 1; Page; 1017pp; English.
XX	
XX	The invention discloses a composition comprising two or more isolated rat
CC	or human polynucleotides or a polynucleotide which represents a fragment,
CC	derivative or allelic variation of the nucleic acid sequence. Also
CC	claimed are a vector comprising the novel polynucleotide, a host cell
CC	comprising the vector, a method for identifying a nucleotide sequence
CC	which is differentially regulated in an animal subjected to pain and a
CC	kit to perform the method, an array, a method for identifying an agent
CC	that increases or decreases the expression of the polynucleotide sequence
CC	that is differentially expressed in neuronal tissue of a first animal
CC	subjected to pain, a method for identifying a compound which regulates
CC	the expression of a polynucleotide sequence which is differentially
CC	expressed in an animal subjected to pain, a method for identifying a
CC	compound that regulates the activity of one or more of the
CC	polynucleotides, a method for producing a pharmaceutical composition, a
CC	method for identifying a compound or small molecule that regulates the
CC	activity in an animal of one or more of the polypeptides given in the
CC	specification, a method for identifying a compound useful in treating
CC	pain and a pharmaceutical composition comprising the one or more
CC	polypeptides or their antibodies. The polynucleotide or the compound that
CC	modulates its activity is useful for preparing a medicament for treating
CC	pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC	injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC	therapy). The sequence presented is a human protein (shown in Table 2 of
CC	the specification) which is differentially expressed during pain. Note:
CC	The sequence data for this patent did not form part of the printed
CC	specification, but was obtained in electronic form directly from WIPO at
CC	ftp.wipo.int/pub/published pct sequences.

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 26, 2005, 12:02:34 ; Search time 19.5347 Seconds
(without alignments)
110.819 Million cell updates/sec

Title: us-10-624-218-2

Perfect score: 29

Sequence: 1 RAAFLKENTALTTRVAELKRVGRCRNI 29

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 513545 seqs, 74649064 residues

Word size : 0

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents_AA.*

- 1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/1/1aa/PCTUS_COMB.pep.*
- 6: /cgn2_6/ptodata/1/1aa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	14	48.3	345	US-09-949-016-10895	Sequence 10895, A
2	9	31.0	335	US-09-949-016-11210	Sequence 11210, A
3	7	24.1	84	US-09-902-540-16607	Sequence 16607, A
4	7	24.1	375	US-09-498-520A-24	Sequence 24, Appl
5	6	20.7	86	US-09-107-532A-4130	Sequence 4130, Ap
6	6	20.7	92	US-09-902-540-13617	Sequence 13617, A
7	6	20.7	105	US-09-583-110-5144	Sequence 5144, Ap
8	6	20.7	135	US-09-107-433-5156	Sequence 5156, Ap
9	6	20.7	141	US-09-489-039A-11896	Sequence 11896, A
10	6	20.7	141	US-09-583-110-3869	Sequence 3869, Ap
11	6	20.7	142	US-09-328-352-6345	Sequence 6345, Ap
12	6	20.7	150	US-09-732-210-626	Sequence 626, App
13	6	20.7	161	US-09-053-197A-27	Sequence 27, Appl
14	6	20.7	161	US-09-085-761A-27	Sequence 27, Appl
15	6	20.7	164	US-09-216-393B-94	Sequence 94, Appl
16	6	20.7	186	US-09-902-540-13724	Sequence 13724, A
17	6	20.7	197	US-09-370-473-12	Sequence 12, Appl
18	6	20.7	199	US-09-252-991A-29327	Sequence 29327, A
19	6	20.7	209	US-09-134-001C-5126	Sequence 5126, Ap
20	6	20.7	220	US-08-297-431B-31	Sequence 31, Appl
21	6	20.7	221	US-08-297-431B-2	Sequence 2, Appl
22	6	20.7	221	US-08-297-431B-4	Sequence 4, Appl
23	6	20.7	221	US-08-297-431B-6	Sequence 6, Appl
24	6	20.7	221	US-08-297-431B-8	Sequence 8, Appl
25	6	20.7	221	US-08-297-431B-10	Sequence 10, Appl
26	6	20.7	221	US-08-297-431B-12	Sequence 12, Appl
27	6	20.7	221	US-08-297-431B-14	Sequence 14, Appl

28	6	20.7	221	3	US-08-297-431B-16	Sequence 16, Appl
29	6	20.7	221	3	US-08-297-431B-18	Sequence 18, Appl
30	6	20.7	221	3	US-08-297-431B-20	Sequence 20, Appl
31	6	20.7	248	4	US-09-902-540-16567	Sequence 16567, A
32	6	20.7	249	4	US-09-134-000C-6781	Sequence 6781, Ap
33	6	20.7	250	4	US-09-583-110-4204	Sequence 4204, Ap
34	6	20.7	255	4	US-09-107-433-2748	Sequence 2748, Ap
35	6	20.7	266	4	US-09-252-991A-24870	Sequence 24870, A
36	6	20.7	267	4	US-09-107-532A-5683	Sequence 5683, Ap
37	6	20.7	281	4	US-09-537-654-4	Sequence 4, Appl
38	6	20.7	294	4	US-09-537-654-6	Sequence 6, Appl
39	6	20.7	294	4	US-09-248-796A-19300	Sequence 19300, A
40	6	20.7	296	4	US-09-270-767-42478	Sequence 42478, A
41	6	20.7	302	4	US-09-270-767-33326	Sequence 33326, A
42	6	20.7	302	4	US-09-270-767-48543	Sequence 48543, A
43	6	20.7	316	4	US-09-949-016-9954	Sequence 9954, Ap
44	6	20.7	316	4	US-08-444-646-6	Sequence 6, Appl
45	6	20.7	355	2	US-08-444-646-6	Sequence 6, Appl

ALIGNMENTS

RESULT 1

US-09-949-016-10895
; Sequence 10895, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10895
; LENGTH: 345
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-10895

Query Match 48.3%; Score 14; DB 4; Length 345;
Best Local Similarity 100.0%; Pred. No. 1.4e-06;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RAAFLKENTALT 14

Db 306 RAAFLKENTALT 319

RESULT 2

US-09-949-016-11210
; Sequence 11210, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498

```

; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11210
; LENGTH: 335
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-11210

Query Match      31.0%; Score 9; DB 4; Length 335;
Best Local Similarity 100.0%; Pred. No. 0.16;
Matches          9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY              1 RAAFLXEN 9
                |||||
Db               296 RAAFLXEN 304

RESULT 3
US-09-902-540-16607
; Sequence 16607, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 16607
; LENGTH: 84
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-16607

Query Match      24.1%; Score 7; DB 4; Length 84;
Best Local Similarity 100.0%; Pred. No. 4.9;
Matches          7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY              1 RAAFLX 7
                |||||
Db               70 RAAFLX 76

RESULT 4
US-09-498-520A-24
; Sequence 24, Application US/09498520A
; Patent No. 6613553
; GENERAL INFORMATION:
; APPLICANT: Rock, Charles O
; APPLICANT: Heath, Richard J
; TITLE OF INVENTION: No. 6613553e1 Enoyl Reductases and Methods of Use Thereof
; FILE REFERENCE: S3-0022
; CURRENT APPLICATION NUMBER: US/09/498,520A
; CURRENT FILING DATE: 2000-02-04
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 24
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis rv1533
US-09-498-520A-24

Query Match      24.1%; Score 7; DB 4; Length 375;
Best Local Similarity 100.0%; Pred. No. 19;
Matches          7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

;

```


; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof

; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 13617
; LENGTH: 92
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-13617

Query Match 20.7%; Score 6; DB 4; Length 92;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 10 TALRTR 15
| | | | |
Db 51 TALRTR 56

RESULT 7

US-09-583-110-5144
; Sequence 5144, Application US/09583110
; Patent No. 6699703
; GENERAL INFORMATION:

; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; FILE REFERENCE: PATH00-07A
; CURRENT APPLICATION NUMBER: US/09/583,110

; CURRENT FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 5144

; LENGTH: 105
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-583-110-5144

Query Match 20.7%; Score 6; DB 4; Length 105;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 ENTALR 13
| | | | |
Db 37 ENTALR 42

RESULT 8

US-09-107-433-5156
; Sequence 5156, Application US/09107433
; Patent No. 6800744
; GENERAL INFORMATION:

; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
; SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE FOR DIAGNO
; THERAPEUTICS

; NUMBER OF SEQUENCES: 5206
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:

; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: <Unknown>
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: <Unknown>
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,433
; FILING DATE: 30-Jun-1998

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/ 085131
; FILING DATE: May 12, 1998
; APPLICATION NUMBER: 60/051553
; FILING DATE: July 2, 1997

; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-011
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277

; INFORMATION FOR SEQ ID NO: 5156:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 135 amino acids
; TYPE: amino acid
; TOPOLOGY: linear

; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pneumoniae
; FEATURE:

; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...135
; SEQUENCE DESCRIPTION: SEQ ID NO: 5156:
US-09-107-433-5156

Query Match 20.7%; Score 6; DB 4; Length 135;
Best Local Similarity 100.0%; Pred. No. 77;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 ENTALR 13
| | | | |
Db 67 ENTALR 72

RESULT 9

US-09-489-039A-11896
; Sequence 11896, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:

; APPLICANT: Gary Bretton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A

; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 11896

; LENGTH: 141
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-11896

Query Match 20.7%; Score 6; DB 4; Length 141;
Best Local Similarity 100.0%; Pred. No. 81;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 12 LRTRVA 17
| | | | |
Db 4 LRTRVA 9

RESULT 10

Thu Jul 28 11:34:41 2005

```

US-09-583-110-3869
; Sequence 3869, Application US/09583110
; Patent No. 6699703
; GENERAL INFORMATION:
; APPLICANT: LYNN Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; FILE REFERENCE: PATH00-07A
; CURRENT APPLICATION NUMBER: US/09/583,110
; CURRENT FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 3869
; LENGTH: 141
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-583-110-3869

Query Match      20.7%; Score 6; DB 4; Length 141;
Best Local Similarity 100.0%; Pred. No. 81;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      4 FLEKEN 9
      |||||
Db      21 FLEKEN 26

RESULT 11
US-09-328-352-6345
; Sequence 6345, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 6345
; LENGTH: 142
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-6345

Query Match      20.7%; Score 6; DB 4; Length 142;
Best Local Similarity 100.0%; Pred. No. 81;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 AFLEKE 8
      |||||
Db      14 AFLEKE 19

RESULT 12
US-09-732-210-626
; Sequence 626, Application US/09732210
; Patent No. 6573361
; GENERAL INFORMATION:
; APPLICANT: Bunkers, Greg J.
; APPLICANT: Liang, Jihong
; APPLICANT: Mittanck, Cindy A.
; APPLICANT: Seale, Jeffrey W.
; APPLICANT: Wu, Yonnie S.
; TITLE OF INVENTION: Anti-fungal Proteins and Methods for Their Use
; FILE REFERENCE: 38-21(15036)B
; CURRENT APPLICATION NUMBER: US/09/732,210
; CURRENT FILING DATE: 2000-12-07

```

```

; PRIOR APPLICATION NUMBER: US 60/169,513
; PRIOR FILING DATE: 1999-12-07
; PRIOR APPLICATION NUMBER: US 60/169,340
; PRIOR FILING DATE: 1999-12-07
; NUMBER OF SEQ ID NOS: 1753
; SEQ ID NO 626
; LENGTH: 150
; TYPE: PRT
; ORGANISM: Thermus aquaticus (subsp. thermophilus)
US-09-732-210-626

Query Match      20.7%; Score 6; DB 4; Length 150;
Best Local Similarity 100.0%; Pred. No. 85;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      20 RKRVR 25
      |||||
Db      16 RKRVR 21

RESULT 13
US-09-053-197A-27
; Sequence 27, Application US/09053197A
; Patent No. 6022952
; GENERAL INFORMATION:
; APPLICANT: Weiner, Joel H.
; APPLICANT: Turner, Raymond J.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR PROTEIN
; TITLE OF INVENTION: SECRETION
; NUMBER OF SEQUENCES: 72
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/09/053,197A
; APPLICATION NUMBER: US/09/053,197A
; FILING DATE: 01-APR-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MacKnight, Kamrin T.
; REGISTRATION NUMBER: 38,230
; REFERENCE/DOCKET NUMBER: UALB-03293
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 161 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-09-053-197A-27

Query Match      20.7%; Score 6; DB 3; Length 161;
Best Local Similarity 100.0%; Pred. No. 91;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      17 AELRKR 22
      |||||
Db      5 AELRKR 10

RESULT 14
US-09-085-761A-27

```

; Sequence 27, Application US/09085761A
; Patent No. 6335178
; GENERAL INFORMATION:
; APPLICANT: Weiner, Joel H.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR PROTEIN
; TITLE OF INVENTION: SECRETION
; NUMBER OF SEQUENCES: 77
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/085,761A
; FILING DATE: 28-MAY-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Cartoll, Peter G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: UALB-03356
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 161 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-09-085-761A-27

Query Match 20.7%; Score 6; DB 3; Length 161;
Best Local Similarity 100.0%; Pred. No. 91;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 17 AELRKR 22
| | | | |
Db 5 AELRKR 10

RESULT 15
US-09-216-393B-94
; Sequence 94, Application US/09216393B
; Patent No. 6514694
; GENERAL INFORMATION:
; APPLICANT: Milhausen, Michael James
; TITLE OF INVENTION: TOKOPLASMA GONDII PROTEINS, NUCLEIC ACID MOLECULES, AND USES THEREOF
; FILE REFERENCE: TX-1-C2
; CURRENT APPLICATION NUMBER: US/09/216,393B
; CURRENT FILING DATE: 1998-12-18
; PRIOR APPLICATION NUMBER: 08/994,825
; PRIOR FILING DATE: 1997-12-19
; NUMBER OF SEQ ID NOS: 366
; SOFTWARE: Patent In version 3.1
; SEQ ID NO 94
; LENGTH: 164
; TYPE: PRT
; ORGANISM: Toxoplasma gondii
; US-09-216-393B-94

Query Match 20.7%; Score 6; DB 4; Length 164;
Best Local Similarity 100.0%; Pred. No. 92;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 4 FLEKEN 9
| | | | |
Db 139 FLEKEN 144

Search completed: July 26, 2005, 12:22:30
Job time : 19.5347 secs

This page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 26, 2005, 12:08:05 ; Search time 68.2708 Seconds
(without alignments)
165.236 Million cell updates/sec

Title: US-10-624-218-2

Perfect score: 29

Sequence: 1 RAAFLEKENTALTRVAELRKVGRCRNI 29

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 1741741 seqs, 388992284 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1741741

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/US06_PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09D_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US10F_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/2/pubpaa/US11A_PUBCOMB.pep.*
- 20: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
- 21: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 22: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	29	100.0	29	16	US-10-624-218-2
2	14	48.3	29	16	US-10-624-218-1
3	14	48.3	303	16	US-10-408-765A-507
4	8	27.6	498	16	US-10-781-014-564
5	8	27.6	675	9	US-09-738-626-5917
6	8	27.6	675	17	US-10-494-836-116
7	7	24.1	43	16	US-10-624-218-3
8	7	24.1	184	15	US-10-424-599-173473
9	7	24.1	249	15	US-10-282-122A-52082
10	7	24.1	277	16	US-10-437-963-130893
11	7	24.1	350	16	US-10-437-963-108061

12	7	24.1	386	16	US-10-437-963-143779	Sequence 143779,
13	7	24.1	401	14	US-10-214-473-66	Sequence 66, Appl
14	7	24.1	401	14	US-10-272-490-66	Sequence 66, Appl
15	7	24.1	404	16	US-10-425-115-255165	Sequence 255165,
16	7	24.1	515	16	US-10-437-963-141509	Sequence 141509,
17	7	24.1	767	15	US-10-149-310-90	Sequence 90, Appl
18	7	24.1	790	15	US-10-369-493-21510	Sequence 21510, A
19	7	24.1	878	16	US-10-437-963-157596	Sequence 157596,
20	6	20.7	35	13	US-10-029-217A-20	Sequence 20, Appl
21	6	20.7	40	15	US-10-424-599-255314	Sequence 255314,
22	6	20.7	43	16	US-10-624-218-4	Sequence 4, Appl
23	6	20.7	54	16	US-10-437-963-149860	Sequence 149860,
24	6	20.7	62	16	US-10-767-701-48001	Sequence 48001, A
25	6	20.7	66	16	US-10-425-115-343881	Sequence 343881,
26	6	20.7	69	16	US-10-437-963-108561	Sequence 108561,
27	6	20.7	76	16	US-10-425-115-241929	Sequence 241929,
28	6	20.7	76	16	US-10-425-115-367583	Sequence 367583,
29	6	20.7	79	15	US-10-424-599-255597	Sequence 255597,
30	6	20.7	81	15	US-10-424-599-151698	Sequence 151698,
31	6	20.7	93	16	US-10-425-115-300995	Sequence 300995,
32	6	20.7	95	14	US-10-156-761-12442	Sequence 12442, A
33	6	20.7	98	17	US-10-732-923-8065	Sequence 8065, Ap
34	6	20.7	107	16	US-10-472-928-1832	Sequence 1832, Ap
35	6	20.7	107	16	US-10-425-115-192433	Sequence 192433,
36	6	20.7	114	10	US-09-769-744A-132	Sequence 132, App
37	6	20.7	114	11	US-08-884-408A-5580	Sequence 5580, Ap
38	6	20.7	115	15	US-10-424-599-154545	Sequence 154545,
39	6	20.7	119	15	US-10-335-977-7311	Sequence 7311, Ap
40	6	20.7	120	16	US-10-425-115-331682	Sequence 331682,
41	6	20.7	128	17	US-10-732-923-7136	Sequence 7136, Ap
42	6	20.7	130	15	US-10-335-977-7313	Sequence 7313, Ap
43	6	20.7	131	15	US-10-335-977-7312	Sequence 7312, Ap
44	6	20.7	135	16	US-10-437-963-119669	Sequence 119669,
45	6	20.7	135	16	US-10-425-115-220155	Sequence 220155,

ALIGNMENTS

RESULT 1

US-10-624-218-2
; Sequence 2, Application US/10624218
; Publication No. US20040171158A1
; GENERAL INFORMATION:
; APPLICANT: Mikheeva, Galina
; TITLE OF INVENTION: Adenoviral Vector Incorporating Zipper
; TITLE OF INVENTION: Peptide-Modified Fiber Protein and Uses Thereof
; FILE REFERENCE: D6463
; CURRENT APPLICATION NUMBER: US/10/624,218
; CURRENT FILING DATE: 2003-07-22
; PRIOR APPLICATION NUMBER: US 60/397,951
; PRIOR FILING DATE: 2002-07-22
; NUMBER OF SEQ ID NOS: 4
; SEQ ID NO 2
; LENGTH: 29
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: PEPTIDE
; OTHER INFORMATION: zipper peptide R R34
US-10-624-218-2

Query Match 100.0%; Score 29; DB 16; Length 29;
Best Local Similarity 100.0%; Pred. No. 6e-22;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RAAFLEKENTALTRVAELRKVGRCRNI 29
DB 1 RAAFLEKENTALTRVAELRKVGRCRNI 29

RESULT 2

Thu Jul 28 11:34:41 2005

```
US-10-624-218-1
; Sequence 1, Application US/10624218
; Publication No. US20040171159A1
; GENERAL INFORMATION:
; APPLICANT: Korokhov, Nikolay
; APPLICANT: Mikheeva, Galina
; TITLE OF INVENTION: Adenoviral Vector Incorporating Zipper
; TITLE OF INVENTION: Peptide-Modified Fiber Protein and Uses Thereof
; FILE REFERENCE: D6463
; CURRENT APPLICATION NUMBER: US/10/624,218
; CURRENT FILING DATE: 2003-07-22
; PRIOR APPLICATION NUMBER: US 60/397,951
; PRIOR FILING DATE: 2002-07-22
; NUMBER OF SEQ ID NOS: 4
; SEQ ID NO 1
; LENGTH: 29
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: PEPTIDE
; OTHER INFORMATION: zipper peptide E E34
US-10-624-218-1
Query Match 48.3%; Score 14; DB 16; Length 29;
Best Local Similarity 100.0%; Pred. No. 7.4e-07;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RAAFLKENTALT 14
Db 1 RAAFLKENTALT 14
|||||

RESULT 3
US-10-408-765A-507
; Sequence 507, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 507
; LENGTH: 303
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-507
Query Match 48.3%; Score 14; DB 16; Length 303;
Best Local Similarity 100.0%; Pred. No. 6e-06;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RAAFLKENTALT 14
Db 264 RAAFLKENTALT 277
|||||

RESULT 4
US-10-781-014-564
; Sequence 564, Application US/10781014
; Publication No. US20040180408A1
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Kroger, Burkhard
```

```
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberhauer, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
; TITLE OF INVENTION: INVOLVED IN CARBON METABOLISM AND ENERGY
; FILE REFERENCE: BGI-126PCPN
; CURRENT APPLICATION NUMBER: US/10/781,014
; CURRENT FILING DATE: 2004-02-17
; PRIOR APPLICATION NUMBER: US 09/602,740
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 60/141,031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: 60/143,208
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 60/151,572
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19931412.8
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931413.6
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931419.5
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931420.9
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931424.1
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931428.4
; PRIOR FILING DATE: 1999-07-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 784
; SEQ ID NO 564
; LENGTH: 498
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-10-781-014-564
Query Match 27.6%; Score 8; DB 16; Length 498;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 14 TRVAELRK 21
Db 297 TRVAELRK 304
|||||

RESULT 5
US-09-738-626-5917
; Sequence 5917, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIO
; APPLICANT: OCHIALI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
```

; SEQ ID NO 5917
; LENGTH: 675
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-5917

Query Match 27.6%; Score 8; DB 9; Length 675;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 TRVAELRK 21
|||||
DB 474 TRVAELRK 481

RESULT 6

US-10-494-836-116
; Sequence 116, Application US/10494836
; Publication No. US20050014233A1
; GENERAL INFORMATION:
; APPLICANT: Zelder, Oskar
; APPLICANT: Pompejus, Markus
; APPLICANT: Schroder, Hartwig
; APPLICANT: Kroger, Burkhard
; APPLICANT: Klopffrogge, Corinna
; APPLICANT: Haberhauer, Gregor
; TITLE OF INVENTION: Genes coding for proteins of carbon metabolism and energy product
; FILE REFERENCE: BG1-167US
; CURRENT APPLICATION NUMBER: US/10/494,836
; CURRENT FILING DATE: 2004-05-05
; PRIOR APPLICATION NUMBER: PCT/EP02/12135
; PRIOR FILING DATE: 2002-10-31
; PRIOR APPLICATION NUMBER: DE 101 54 270.4
; PRIOR FILING DATE: 2001-11-05
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 116
; LENGTH: 675
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-10-494-836-116

Query Match 27.6%; Score 8; DB 17; Length 675;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 TRVAELRK 21
|||||
DB 474 TRVAELRK 481

RESULT 7

US-10-624-218-3
; Sequence 3, Application US/10624218
; Publication No. US20040171158A1
; GENERAL INFORMATION:
; APPLICANT: Korokhov, Nikolay
; APPLICANT: Mikheeva, Galina
; TITLE OF INVENTION: Adenoviral Vector Incorporating Zipper
; FILE REFERENCE: D6463
; CURRENT APPLICATION NUMBER: US/10/624,218
; CURRENT FILING DATE: 2003-07-22
; PRIOR APPLICATION NUMBER: US 60/397,951
; PRIOR FILING DATE: 2002-07-22
; NUMBER OF SEQ ID NOS: 4
; SEQ ID NO 3
; LENGTH: 43
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: PEPTIDE
; OTHER INFORMATION: zipper peptide EE12RR345L
US-10-624-218-3

Query Match 24.1%; Score 7; DB 16; Length 43;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 TRVAELRK 20
|||||
DB 17 TRVAELRK 23

RESULT 8

US-10-424-599-173473
; Sequence 173473, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 173473
; LENGTH: 184
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_127664C.1.pep
US-10-424-599-173473

Query Match 24.1%; Score 7; DB 15; Length 184;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 RVAELRK 21
|||||
DB 52 RVAELRK 58

RESULT 9

US-10-282-122A-52082
; Sequence 52082, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELTRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09

```

; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 52082
; LENGTH: 249
; TYPE: PRT
; ORGANISM: Clostridium acetobutylicum
US-10-282-122A-52082

Query Match      24.1%; Score 7; DB 15; Length 249;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      18 ELRRKVG 24
Db      80 ELRRKVG 86
      |||||

RESULT 10
US-10-437-963-130893
; Sequence 130893, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 130893
; LENGTH: 277
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_33011C.1.pep
US-10-437-963-130893

Query Match      24.1%; Score 7; DB 16; Length 277;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      16 VAE LRKR 22
Db      220 VAE LRKR 226
      |||||

RESULT 11
US-10-437-963-108061
; Sequence 108061, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei

```

```

; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 108061
; LENGTH: 350
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_12351C.1.pep
US-10-437-963-108061

Query Match      24.1%; Score 7; DB 16; Length 350;
Best Local Similarity 100.0%; Pred. No. 74;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      3 AFLEKEN 9
Db      343 AFLEKEN 349
      |||||

RESULT 12
US-10-437-963-143779
; Sequence 143779, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 143779
; LENGTH: 386
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_44655C.1.pep
US-10-437-963-143779

Query Match      24.1%; Score 7; DB 16; Length 386;
Best Local Similarity 100.0%; Pred. No. 81;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      11 ALRTRVA 17
Db      19 ALRTRVA 25
      |||||

RESULT 13
US-10-214-473-66
; Sequence 66, Application US/10214473
; Publication No. US2003014843A1
; GENERAL INFORMATION:
; APPLICANT: Zhao, Lishan
; APPLICANT: Mathur, Eric J.
; APPLICANT: Weiner, David
; APPLICANT: Richardson, Toby
; APPLICANT: Milan, Aileen

```



```

; APPLICANT: Burk, Mark J.
; APPLICANT: Han, Bin
; APPLICANT: Short, Jay M.
; TITLE OF INVENTION: EPOXIDE HYDROLASES, NUCLEIC ACIDS ENCODING
; TITLE OF INVENTION: THEM AND METHODS OF MAKING AND USING THEM
; FILE REFERENCE: 09010-600001
; CURRENT APPLICATION NUMBER: US/10/214,473
; CURRENT FILING DATE: 2002-08-05
; PRIOR APPLICATION NUMBER: US 60/309,478
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US 60/393,378
; PRIOR FILING DATE: 2002-07-03
; NUMBER OF SEQ ID NOS: 83
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 66
; LENGTH: 401
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Obtained from an environmental source
US-10-214-473-66

```

```

Query Match      24.1%; Score 7; DB 14; Length 401;
Best Local Similarity 100.0%; Pred. No. 84;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      17 AELRKRV 23
        |||||
DB      28 AELRKRV 34

```

```

RESULT 14
US-10-272-490-66
; Sequence 66, Application US/10272490
; Publication No. US20030148490A1
; GENERAL INFORMATION:
; APPLICANT: Zhao, Lishan
; APPLICANT: Mathur, Eric J.
; APPLICANT: Weiner, David
; APPLICANT: Richardson, Toby
; APPLICANT: Milan, Aileen
; APPLICANT: Burk, Mark J.
; APPLICANT: Han, Bin
; APPLICANT: Short, Jay M.
; TITLE OF INVENTION: EPOXIDE HYDROLASES, NUCLEIC ACIDS ENCODING THEM AND METHODS
; TITLE OF INVENTION: OF MAKING AND USING THEM
; FILE REFERENCE: 09010-831001
; CURRENT APPLICATION NUMBER: US/10/272,490
; CURRENT FILING DATE: 2002-10-10
; PRIOR APPLICATION NUMBER: US 10/214,473
; PRIOR FILING DATE: 2002-08-05
; PRIOR APPLICATION NUMBER: US 60/309,478
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US 60/393,378
; PRIOR FILING DATE: 2002-07-03
; NUMBER OF SEQ ID NOS: 94
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 66
; LENGTH: 401
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Obtained from an environmental source
US-10-272-490-66

```

```

Query Match      24.1%; Score 7; DB 14; Length 401;
Best Local Similarity 100.0%; Pred. No. 84;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      17 AELRKRV 23
        |||||
DB      28 AELRKRV 34

```

```

RESULT 15
US-10-425-115-255165
; Sequence 255165, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 255165
; LENGTH: 404
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_164289C.1.pep
US-10-425-115-255165

Query Match      24.1%; Score 7; DB 16; Length 404;
Best Local Similarity 100.0%; Pred. No. 84;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      16 VAE LRKR 22
        |||||
DB      250 VAE LRKR 256

```

```

Search completed: July 26, 2005, 12:28:15
Job time : 68.2708 secs

```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 26, 2005, 12:00:54 ; Search time 13.2917 Seconds
(without alignments)
209.927 Million cell updates/sec

Title: US-10-624-218-2

Perfect score: 29

Sequence: 1 RAAPLEKENTALRTTRVAELKRKVGRCRNI 29

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 283416 seqs, 96216763 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : PIR 79:**

1: pir1:**

2: pir2:**

3: pir3:**

4: pir4:**

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	14	48.3	261	2 A40579	trans-activating t
2	14	48.3	293	2 S50109	vitellogenin gene-
3	14	48.3	303	2 G02360	thyrotroph embryon
4	14	48.3	313	2 A41524	transcription fact
5	9	31.0	325	2 A55558	albumin D-box bind
6	8	27.6	216	1 C69129	riboflavin-specifi
7	7	24.1	73	1 DNBPD8	DNA-binding protei
8	7	24.1	153	2 T40951	40s ribosomal prot
9	7	24.1	154	2 T37489	40s ribosomal prot
10	7	24.1	205	2 S08867	crystallin (clone
11	7	24.1	219	2 T45997	hypothetical prote
12	7	24.1	222	2 B84410	hypothetical prote
13	7	24.1	249	2 G97110	ATPase component o
14	7	24.1	252	2 C64426	phosphate transport
15	7	24.1	320	2 AF2714	conserved hypothet
16	7	24.1	352	2 C97496	hypothetical prote
17	7	24.1	364	2 D75532	periplasmic serine
18	7	24.1	375	2 B70500	hypothetical prote
19	7	24.1	392	1 S74675	hypothetical prote
20	7	24.1	767	2 T41344	probable zinc-fing
21	7	24.1	773	2 D75031	hydrogenase expres
22	7	24.1	790	1 G69071	heavy-metal-transp
23	7	24.1	845	2 I48176	synaptonemal compl
24	7	24.1	1070	2 T06733	kinesin homolog f2
25	6	20.7	57	2 AB0724	conserved hypothet
26	6	20.7	86	2 S41060	probable secretory
27	6	20.7	95	2 A75534	ribosomal protein
28	6	20.7	105	2 D95108	conserved hypothet
29	6	20.7	114	2 E97976	conserved hypothet

30	6	20.7	115	2 B75105	hypothetical prote
31	6	20.7	118	2 B85821	unknown protein en
32	6	20.7	118	2 G50973	hypothetical prote
33	6	20.7	121	2 T11787	probable protein t
34	6	20.7	130	2 A87666	death on curing pr
35	6	20.7	136	2 AH3394	lactoylglutathione
36	6	20.7	136	2 D71029	hypothetical prote
37	6	20.7	138	2 T49183	hypothetical prote
38	6	20.7	141	2 E85138	galactose-6-phosph
39	6	20.7	141	2 C98006	galactose-6-phosph
40	6	20.7	147	2 A72248	ribosomal protein
41	6	20.7	149	2 A69948	phage-related prot
42	6	20.7	150	2 C71939	phage-related prot
43	6	20.7	150	2 C64648	ferric uptake regu
44	6	20.7	151	1 B93375	phosphohistidine p
45	6	20.7	151	2 E95378	Syrb regulatory pr

ALIGNMENTS

RESULT 1

A40579 trans-activating transcription regulator TEF.- rat

C:Species: Rattus norvegicus (Norway rat)

C>Date: 05-Jun-1992 #sequence_revision 05-Jun-1992 #text_change 09-Jul-2004

C:Accession: A40579

R:Drölet, D.W.; Scully, K.M.; Simmons, D.M.; Wegner, M.; Chu, K.; Swanson, L.W.; Rosenf

Genes Dev. 5, 1739-1753, 1991

A:Title: TEF, a transcription factor expressed specifically in the anterior pituitary d

A:Reference number: A40579; MUID:92009166; PMID:1916262

A:Accession: A40579

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-261 <DR>

A:Cross-references: UNIPROT:P41224; GB:S58745; NID:G237084; PIDN:AA820032.1; PID:G23708

C:Keywords: DNA binding; transcription regulation

Query Match 48.3%; Score 14; DB 2; Length 261;

Best Local Similarity 100.0%; Pred. No. 9.3e-07;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RAAPLEKENTALRT 14

DB 222 RAAPLEKENTALRT 235

RESULT 2

S50109 vitellogenin gene-binding protein VBP, beta/beta isoform - chicken

C:Species: Gallus gallus (chicken)

C>Date: 14-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 09-Jul-2004

C:Accession: S50109; S50111

R:Burch, J.B.E.; Davis, D.L.

Nucleic Acids Res. 22, 4733-4741, 1994

A:Title: Alternative promoter usage and splicing options result in the differential exp-

ors.

A:Reference number: S50109; MUID:95075656; PMID:7984425

A:Accession: S50109

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-293 <BUR>

A:Cross-references: EMBL:U09223; UNIPROT:Q92172; EMBL:U09221; NID:G483937; PIDN:AAA82156.1; PID:G483

A:Accession: S50111

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-34 <EU2>

A:Cross-references: EMBL:U09223; NID:G483823; PIDN:AAA82158.1; PID:G483824

C:Keywords: transcription factor

Query Match 48.3%; Score 14; DB 2; Length 293;

Best Local Similarity 100.0%; Pred. No. 1e-06;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;


```
RESULT 7
DNBPDB
DNA-binding protein Ner - phase D108
C:Species: phase D108
C:Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 09-Jul-2004
C:Accession: A24680; S07931
R:Tollas, P.P.; DuBow, M.S.
EMBO J. 4, 3031-3037, 1985
A:Title: The cloning and characterization of the bacteriophage D108 regulatory DNA-binding
A:Reference number: A24680; MUID:86055744; PMID:2998774
A:Accession: A24680
A:Molecule type: DNA
A:Residues: 1-73 <POL>
A:Cross-references: UNIPROT:P06903; GB:M26291; NID:G166194; PIDN:AAA32206.1; PID:G166195
R:Mizuuchi, M.; Weisberg, R.A.; Mizuuchi, K.
Nucleic Acids Res. 14, 3813-3825, 1986
A:Title: DNA sequence of the control region of phage D108: the N-terminal amino acid seq
A:Reference number: S07370; MUID:86232621; PMID:3012481
A:Accession: S07931
A:Molecule type: DNA
A:Residues: 1-73 <MIZ>
A:Cross-references: EMBL:X03847; NID:G15439; PIDN:CAA27475.1; PID:G15441
C:Genetics:
A:Gene: ner
C:Superfamily: phage D108 DNA-binding protein
C:Keywords: DNA binding

Query Match      24.1%; Score 7; DB 1; Length 73;
Best Local Similarity 100.0%; Pred. No. 3.4;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 VAE LRKR 22
    |||||
DB 18 VAE LRKR 24

RESULT 8
T40951
40s ribosomal protein s15 - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: T40951
R:Wood, V.; Rajandream, M.A.; Barrell, B.G.; Volckaert, G.
submitted to the EMBL Data Library, February 1999
A:Reference number: Z21940
A:Accession: T40951
A:Status: preliminary; translated from GB/EMBL/DBBJ
A:Molecule type: DNA
A:Residues: 1-153 <WOO>
A:Cross-references: UNIPROT:Q94715; EMBL:AL035592; PIDN:CAB38159.1; GSPDB:GNO00068; SPDB:
A:Experimental source: strain 972h-; cosmid c1393
C:Genetics:
A:Gene: SPDB:SPCC1393.03
A:Map position: 3
C:Superfamily: ribosomal protein S19/S15

Query Match      24.1%; Score 7; DB 2; Length 153;
Best Local Similarity 100.0%; Pred. No. 6.6;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 RVAELRKR 21
    |||||
DB 10 RVAELRKR 16

RESULT 9
T37489
40s ribosomal protein s15 - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: T37489
R:Bothe, G.; Pohl, T.; McDougall, R.C.; Rajandream, M.A.; Barrell, B.G.
```

```
submitted to the EMBL Data Library, November 1999
A:Reference number: Z21718
A:Accession: T37489
A:Status: preliminary; translated from GB/EMBL/DBBJ
A:Molecule type: DNA
A:Residues: 1-154 <BOT>
A:Cross-references: UNIPROT:Q9UTQ6; EMBL:AL132769; PIDN:CAB59883.1; GSPDB:GNO00066; SPDB:
A:Experimental source: strain 972h-; cosmid c1071
C:Genetics:
A:Gene: SPDB:SPAC1071.07c
A:Map position: 1
C:Superfamily: ribosomal protein S19/S15

Query Match      24.1%; Score 7; DB 2; Length 154;
Best Local Similarity 100.0%; Pred. No. 6.6;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 RVAELRKR 21
    |||||
DB 11 RVAELRKR 17

RESULT 10
S06867
crystallin (clone pS11) - Sloane's squid
N:Alternate names: glutathione transferase homolog
C:Species: Ommastrephes sloanei pacificus (Sloane's squid)
C:Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 09-Jul-2004
C:Accession: S06867
R:Tomarev, S.I.; Zinovieva, R.D.
Nature 336, 86-88, 1988
A:Title: Squid major lens polypeptides are homologous to glutathione S-transferases sub
A:Reference number: S06442; MUID:89040202; PMID:3185725
A:Accession: S06867
A:Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-205 <TOM>
A:Cross-references: UNIPROT:P18426; GB:M36938; NID:G159855; PIDN:AAA63411.1; PID:G15985
C:Superfamily: glutathione transferase
C:Keywords: dimer; eye lens

Query Match      24.1%; Score 7; DB 2; Length 205;
Best Local Similarity 100.0%; Pred. No. 8.5;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 ALRTRVA 17
    |||||
DB 182 ALRTRVA 188

RESULT 11
T45997
hypothetical protein F9D24.280 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 09-Jul-2004
C:Accession: T45997
R:D'Angelo, M.; Vezzi, A.; Modesto, D.; Pigazzi, M.; Valle, G.; Mewes, H.W.; Lemcke, K.
submitted to the Protein Sequence Database, January 2000
A:Reference number: Z23011
A:Accession: T45997
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-219 <DAN>
A:Cross-references: UNIPROT:Q9M2H9; EMBL:AL137081
A:Experimental source: cultivar Columbia; BAC clone F9D24
C:Genetics:
A:Map position: 3
A:Introns: 85/3
A:Note: F9D24.280
C:Superfamily: Arabidopsis thaliana hypothetical protein F9D24.210

Query Match      24.1%; Score 7; DB 2; Length 219;
Best Local Similarity 100.0%; Pred. No. 9;
```

Matches		7;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
QY	3	AFLEKEN	9							
Db	197	AFLEKEN	203							
RESULT 12										
B84410										
hypothetical protein Vng2607c [imported] - Halobacterium sp. NRC-1										
C;Species: Halobacterium sp. NRC-1										
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004										
C;Accession: B84410										
R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Laaky, S.; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jabilg, Jung, K.H.; Alam, M.; Freitas, T.										
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000										
A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li										
A;Title: Genome sequence of Halobacterium species NRC-1.										
A;Reference number: A84160; MUID:20504483; PMID:11016950										
A;Accession: B84410										
A;Status: preliminary										
A;Molecule type: DNA										
A;Residues: 1-222 <STO>										
A;Cross-references: UNIPROT:Q9HWC5; GB:AE004437; NID:gl0581987; PIDN:AAG20646.1; GSPDB:G										
C;Genetics:										
A;Gene: VNG2607C										
C;Superfamily: conserved hypothetical protein MJ1501										
Query Match										
Best Local Similarity										
Matches										
24.1%; Score 7; DB 2; Length 222;										
100.0%; Pred. No. 9.1;										
0; Mismatches 0; Indels 0; Gaps 0;										
QY	13	RTRVAEL	19							
Db	170	RTRVAEL	176							
RESULT 13										
G97110										
ATPase component of ABC-type phosphate transport system CAC1708 [imported] - Clostridium										
C;Species: Clostridium acetobutylicum										
C;Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 09-Jul-2004										
C;Accession: G97110										
R;Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,										
; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.										
J. Bacteriol. 183, 4823-4838, 2001										
A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo										
A;Reference number: A96900; MUID:21359325; PMID:21359325										
A;Accession: G97110										
A;Status: preliminary										
A;Molecule type: DNA										
A;Residues: 1-249 <KUR>										
A;Cross-references: UNIPROT:Q971B0; GB:AE001437; PIDN:AAK79674.1; PID:g15024673; GSPDB:G										
A;Experimental source: Clostridium acetobutylicum ATCC824										
C;Genetics:										
A;Gene: CAC1708										
C;Superfamily: inner membrane protein malk; ATP-binding cassette homology										
Query Match										
Best Local Similarity										
Matches										
24.1%; Score 7; DB 2; Length 249;										
100.0%; Pred. No. 10;										
0; Mismatches 0; Indels 0; Gaps 0;										
QY	18	ELRKRVG	24							
Db	80	ELRKRVG	86							
RESULT 14										
C64426										
phosphate transport system ATP-binding protein - Methanococcus jannaschii										
C;Species: Methanococcus jannaschii										
C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 09-Jul-2004										

C;Accession: C64426		R;Sult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake, ,			
; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.;					
reison, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.					
Science 273, 1058-1073, 1996					
A;Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C					
A;Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii					
A;Reference number: A64300; MUID:96337999; PMID:8688087					
A;Accession: C64426					
A;Status: preliminary; nucleic acid sequence not shown; translation not shown					
A;Molecule type: DNA					
A;Residues: 1-252 <BUL>					
A;Cross-references: UNIPROT:Q58418; GB:U67544; GB:L77117; NID:gl591671; PIDN:AAB99016.1;					
C;Genetics:					
A;Map position: REV942687-941929					
C;Superfamily: inner membrane protein malk; ATP-binding cassette homology					
C;Keywords: ATP; nucleotide binding; P-loop					
F;21-223/Domain: ATP-binding cassette homology <ABC>					
F;38-45/Region: nucleotide-binding motif A (P-loop)					
Query Match		24.1%; Score 7; DB 2; Length 252;			
Best Local Similarity		100.0%; Pred. No. 10;			
Matches		7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	18	ELRKRVG	24		
Db	83	ELRKRVG	89		
RESULT 15					
AF2714					
conserved hypothetical protein Atu1119 [imported] - Agrobacterium tumefaciens (strain C5					
C;Species: Agrobacterium tumefaciens					
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004					
C;Accession: AF2714					
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L					
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayvin, T.; Levy, R.; Li, M.; McClellan					
; Karp, P.; Romero, P.; Zhang, S.					
Science 294, 2317-2323, 2001					
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, I					
ster, E.W.					
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.					
A;Reference number: AB2577; MUID:21608550; PMID:11743193					
A;Accession: AF2714					
A;Status: preliminary					
A;Molecule type: DNA					
A;Residues: 1-320 <KUR>					
A;Cross-references: UNIPROT:Q8UGB9; GB:AE008688; PIDN:AAL42132.1; PID:g17739517; GSPDB:G					
A;Experimental source: strain C58 (Dupont)					
C;Genetics:					
A;Gene: Atu1119					
A;Map position: circular chromosome					
Query Match					
Best Local Similarity					
Matches					
24.1%; Score 7; DB 2; Length 320;					
100.0%; Pred. No. 13;					
0; Mismatches 0; Indels 0; Gaps 0;					
QY	1	RAAFLEX	7		
Db	153	RAAFLEX	159		
Search completed: July 26, 2005, 12:20:49					
Job time : 14.2917 secs					

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 26, 2005, 11:59:44 ; Search time 65.6528 Seconds
(without alignments)
226.194 Million cell updates/sec

Title: US-10-624-218-2

Perfect score: 29

Sequence: 1 RAAFLKENTALTTRVAELKRVGRCRN 29

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 1612378 seqs, 512079187 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : UniProt_03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	14	48.3	140	2 Q723J7	Q723J7 homo sapien
2	14	48.3	221	1 TEF PHOSU	P97516 phodopus su
3	14	48.3	271	1 TEF RAT	P41224 rattus norv
4	14	48.3	273	2 Q726X6	Q726X6 homo sapien
5	14	48.3	293	2 Q92172	Q92172 gallus gall
6	14	48.3	301	1 TEF MOUSE	Q9J1C6 mus musculu
7	14	48.3	303	1 TEF HUMAN	Q10587 homo sapien
8	13	44.8	278	2 Q7SY76	Q7SY76 xenopus lae
9	13	44.8	298	2 Q632P9	Q632P9 xenopus lae
10	9	31.0	188	2 Q6R2I2	Q6R2I2 rattus norv
11	9	31.0	225	2 Q6R2I3	Q6R2I3 rattus norv
12	9	31.0	325	1 DBP HUMAN	Q10586 homo sapien
13	9	31.0	325	1 DBP MOUSE	Q60925 mus musculu
14	9	31.0	325	1 DBP RAT	P16443 rattus norv
15	8	27.6	216	1 R1B7 METTH	Q26337 methanobact
16	8	27.6	313	1 DIMH_MACFA	Q95KJ0 macaca fasc
17	8	27.6	444	2 Q74BQ2	Q74BQ2 geobacter s
18	8	27.6	548	2 Q76EM4	Q76EM4 gluconobact
19	8	27.6	567	2 Q8FNP6	Q8FNP6 corynebacte
20	8	27.6	675	2 Q8NNJ2	Q8NNJ2 corynebacte
21	7	24.1	73	1 NER BPD10	P06903 bacterioph
22	7	24.1	153	1 R1S_KSCHPO	Q94715 schizosacch
23	7	24.1	154	1 R1S8_SCHPO	Q9UCQ6 schizosacch
24	7	24.1	205	1 SC11_OMMSL	P18426 ommastrephe
25	7	24.1	219	2 Q9M2H9	Q9M2H9 arabidopsis
26	7	24.1	222	2 Q9HMC5	Q9HMC5 halobacteri
27	7	24.1	222	2 Q6FPA6	Q6FPA6 candida gla
28	7	24.1	249	1 PSTB_CLOAB	Q971E0 clostridium
29	7	24.1	251	2 Q6LY93	Q6LY93 methanococc
30	7	24.1	252	1 PSTB_METJA	Q58418 methanococc
31	7	24.1	253	2 Q73HU0	Q73HU0 wolbachia p

32 7 24.1 258 1 PSTB_METAC Q8tea8 methanosarc
33 7 24.1 261 2 Q88X00 Q88X00 lactobacill
34 7 24.1 275 2 Q65GX2 Q65GX2 bacillus li
35 7 24.1 278 2 Q882J7 Q882J7 pseudomonas
36 7 24.1 294 2 Q67L16 Q67L16 symbiobacte
37 7 24.1 320 2 Q8UGB9 Q8UGB9 agrobacteri
38 7 24.1 350 2 Q7XSE5 Q7XSE5 oryza sativ
39 7 24.1 352 2 Q7CZY1 Q7CZY1 agrobacteri
40 7 24.1 363 2 Q8KB29 Q8KB29 malvastrum
41 7 24.1 364 2 Q9RXI6 Q9RXI6 deinococcus
42 7 24.1 372 2 Q740U8 Q740U8 mycobacteri
43 7 24.1 375 2 O06179 O06179 mycobacteri
44 7 24.1 375 2 Q7TZX5 Q7TZX5 mycobacteri
45 7 24.1 386 2 Q94DF4 Q94DF4 oryza sativ

ALIGNMENTS

RESULT 1

Q723J7 PRELIMINARY; PRT; 140 AA.
AC Q723J7; TREMBLrel. 25, Created)
DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Hypothetical protein DKFP686D1282 (Fragment).
GN Name=DKFP686D1282;
OS Homo sapiens (Human);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Human retina;
RA Wambutt R., Heubner D., Mewes H.W., Weil B., Amid C., Osanger A.,
Fobo G., Han M., Wiemann S.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the bZIP family.
DR EMBL; BX537848; CAD97856.1; -;
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR004827; TF_bZIP.
DR SMART; SM00338; BRLZ; 1.
DR PROSITE; PS00217; BZIP; 1.
KW DNA-binding; Hypothetical protein; Nuclear protein.
FT NON_TER 1
SQ SEQUENCE 140 AA; 16165 MW; EDD5A96CF193E0BC CRC64;

Query Match 48.3%; Score 14; DB 2; Length 140;
Best Local Similarity 100.0%; Pred.No. 3.9e-06;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RAAFLKENTALT 14
DB 101 RAAFLKENTALT 114

RESULT 2

TEF PHOSU STANDARD; PRT; 221 AA.
ID TEF PHOSU
AC P97516;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Thyrotroph embryonic factor (Fragment).
GN Name=TEF;
OS Phodopus sungorus (Striped hairy-footed hamster) (Djungarian hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Phodopus.
OX NCBI_TaxID=10044;


```

DR EMBL; AL035659; CAB62497.1; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR004827; TF_bZIP.
DR SMART; SM00338; BRL2; 1.
DR PROSITE; PS0217; BZIP; 1.
KW DNA-binding; Nuclear protein.
SQ SEQUENCE 273 AA; 30634 MW; BIA956839CBC4AC4 CRC64;

Query Match 48.3%; Score 14; DB 2; Length 273;
Best Local Similarity 100.0%; Pred. No. 6.9e-06;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RAAFLKENTALT 14
Db 234 RAAFLKENTALT 247

RESULT 5
ID 092172 PRELIMINARY; PRT; 293 AA.
AC 092172;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Vitellonin binding protein (VBP), beta/beta isoform.
GN Name=vbp;
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;
OC Gallus.
OC NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS ALPHA AND BETA), AND ALTERNATIVE PROMOTER
RC STRAIN=White Leghorn;
RX MEDLINE=95075656; PubMed=7984425;
RA Burch J.B., Davis D.L.;
RT "Alternative promoter usage and splicing options result in the
RT differential expression of mRNAs encoding four isoforms of chicken
RT VBP, a member of the PAR subfamily of bZIP transcription factors.";
RL Nucleic Acids Res. 22:4733-4741(1994).
CC -!- SIMILARITY: Belongs to the bZIP family.
DR EMBL; U09221; AAA82156.1; -.
DR PIR; S50109; S50109.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR SMART; SM00338; BRL2; 1.
DR PROSITE; PS0217; BZIP; 1.
KW DNA-binding; Nuclear protein.
SQ SEQUENCE 293 AA; 32736 MW; C5803D510CEE03D6 CRC64;

Query Match 48.3%; Score 14; DB 2; Length 293;
Best Local Similarity 100.0%; Pred. No. 7.4e-06;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RAAFLKENTALT 14
Db 248 RAAFLKENTALT 261

RESULT 6
ID TEF_MOUSE
AC Q9JJC6; Q6QHT6; Q8C610; Q8VD02;
DT 25-OCT-2004 (Rel. 45, Created)
DT 25-OCT-2004 (Rel. 45, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Thyrotroph embryonic factor.
GN Name=TEF;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM ALPHA).
RC STRAIN=BALB/c; TISSUE=Brain;
RX MEDLINE=20261582; PubMed=10799536; DOI=10.1074/jbc.275.19.14524;
RA Krueger D.A., Warner E.A., Dowd D.R.;
RT "Involvement of thyrotroph embryonic factor in calcium-mediated
RT regulation of gene expression.";
RL J. Biol. Chem. 275:14524-14531(2000).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS ALPHA AND BETA), AND ALTERNATIVE PROMOTER
RX USAGE.
RX PubMed=14702338; DOI=10.1074/jbc.M313822200;
RA Zhou J., Hoggatt A.M., Herring B.P.;
RT "Activation of the smooth muscle-specific telokin gene by thyrotroph
RT embryonic factor (TEF).";
RL J. Biol. Chem. 279:15929-15937(2004).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORMS ALPHA AND 2).
RC STRAIN=FVB/N; TISSUE=Breast tumor, and Kidney;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smalios D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [4]
RP SEQUENCE OF 97-196 FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Kidney;
RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nikaide I., Oato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmond S., Gusjcinich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita K.,
RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wyshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of

```


RA Clegg S.M., Copley V.E., Cole C.G., Collier R.E., Connor R.,
RA Conroy D., Corby N.R., Coville G.J., Cox A.V., Davis J., Dawson E.,
RA Dharm D., Dockree C., Dodsorth S.J., Durbin R.M., Ellington A.G.,
RA Evans K.L., Fey J.M., Fleming K., French L., Garner A.A.,
RA Gilbert J.G.R., Goward M.E., Grahm D.V., Griffiths M.N.D., Hall C.,
RA Hall R.E., Hall-Tamlyn G., Heathcott R.W., Ho S., Holmes S.,
RA Hunt S.E., Jones M.C., Kershaw J., Kimberley A.M., King A.,
RA Laird G.K., Langford C.F., Leversha M.A., Lloyd D.M.,
RA Martyn I.D., Mashreghi-Mohammadi M., Matthews L.H., Mccann O.T.,
RA McClay J., McLaren S., McMurray A.A., Milne S.A., Mortimore B.J.C.T.,
RA Odell C.N., Pavitt R., Pearce A.V., Pearson D., Phillimore B.J.C.T.,
RA Phillips S.H., Plumb R.W., Ramsey H., Ramsey Y., Rogers L., Ross M.T.,
RA Scott C.E., Sehra H.K., Skuce C.D., Smalley S., Smith M.L.,
RA Soderlund C., Spraggon L., Stewart C.A., Sulston J.E., Swann R.M.,
RA Vaudin M., Wall M., Wallis J.M., Whiteley M.N., Willey D.L.,
RA Williams L., Williams S., Williamson A., Wilmer T.E., Wilming L.,
RA Wright C.L., Hubbard T., Bentley D.R., Beck S., Rogers J., Shimizu N.,
RA Minoshima S., Kawasaki K., Sasaki T., Asakawa S., Kudoh J.,
RA Shintani A., Shibuya K., Yoshizaki Y., Aoki N., Mitsuyama S.,
RA Roe B.A., Chen F., Chu L., Crabtree J., Deschamps S., Do A., Do T.,
RA Dorman A., Fang F., Fu Y., Hu P., Hua A., Kenton S., Lai H., Lao H.I.,
RA Lewis J., Lewis S., Lin S.-P., Loh P., Mala J.E., Nguyen T., Pan H.,
RA Phan S., Qi S., Qian Y., Ray L., Ren Q., Shaull S., Sloan D., Song L.,
RA Wang Q., Wang Y., Wang Z., White J., Willingham D., Wu H., Yao Z.,
RA Zhan M., Zhang G., Chisoe S., Murray J., Miller N., Minx P.,
RA Fulton R., Johnson D., Bemis G., Bentley D., Bradshaw H., Bourne S.,
RA Cordes M., Du Z., Fulton L., Goela D., Graves T., Hawkins J.,
RA Hinds K., Kemp K., Latreille P., Layman D., Ozersky P., Rohlfing T.,
RA Scheet P., Walker C., Wamsley A., Wohlmann P., Pepin K., Nelson J.,
RA Korf I., Bedell J.A., Hillier L.M., Mardis E., Waterston R.,
RA Wilson R., Emanuel B.S., Shaikh T., Kurahashi H., Saitta S.,
RA Budarf M.L., McDermid H.E., Johnson A., Wong A.C.C., Morrow B.E.,
RA Edelman L., Kim U.J., Shizuya H., Simon M.I., Dumanski J.P.,
RA Peyrard M., Kedra D., Seroussi E., Fransson I., Tapia I., Bruder C.E.,
RA O'Brien K.P., Wilkinson P., Bodenteich A., Hartman K., Hu X.,
RA Khan A.S., Lane L., Tilahun Y., Wright H.;
RA "The DNA sequence of human chromosome 22.";
RA Nature 402:489-495(1999).
RA [6]
RA SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.P., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Raha S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahay J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whitling M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalilus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RA "Generation and initial analysis of more than 15,000 full-length human
RA and mouse cDNA sequences.";
RA Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: Transcription factor that binds to and transactivates
CC the TSHB promoter. Binds to a minimal DNA-binding sequence 5'-
CC [TC]AG[AG]TTA[TC]AG[AG]-3'.
CC -!- SUBUNIT: Binds DNA as a homodimer or a heterodimer. Can form a
CC heterodimer with DBP.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- INDUCTION: Accumulates according to a robust circadian rhythm (By
CC similarity).
CC -!- SIMILARITY: Belongs to the bZIP family. PAR subfamily.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U06935; AAA81373.1; ALT_INIT.
DR EMBL; U40059; AAB06497.1; -.
DR EMBL; CR456592; CAG30478.1; -.
DR EMBL; CR541827; CAG46626.1; -.
DR EMBL; AL035659; CAB62498.1; ALT_INIT.
DR EMBL; BC039258; AAH39258.1; -.
DR EMBL; BC042476; AAH42476.1; -.
DR PIR; G02360; G02360.
DR TRANSFAC; T04876; -.
DR Genew; HGNC:11722; TEF.
DR MIM; 188595; -.
DR GO; GO:0003702; F:RNA polymerase II transcription factor acti. . .; TAS.
DR GO; GO:0006357; P:regulation of transcription from Pol II pro. . .; TAS.
DR InterPro; IPR004827; TF_BZIP.
DR Pfam; PF00170; bZIP; 1.
DR SMART; SM00338; BRLZ; 1.
DR PROSITE; PS0217; bZIP; 1.
DR PROSITE; PS00036; bZIP_BASIC; FALSE_NEG.
DR Activator; Biological rhythms; DNA-binding; Nuclear protein;
KW Transcription regulation.
FT DOMAIN 166 217 Pro-rich (Proline/Acidic region (PAR)).
FT DNA BIND 239 257 Basic motif.
FT DOMAIN 268 282 Leucine-zipper.
FT CONFLICT 54 54 K -> E (in Ref. 1).
FT SEQUENCE 303 AA; 33247 MW; 4A87B7BFA7248C6F CRC64;

Query Match 48.3%; Score 14; DB 1; Length 303;
Best Local Similarity 100.0%; Pred. No. 7.6e-06;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RAAFLKENTALRT 14
Db 264 RAAFLKENTALRT 277

RESULT 8
Q7SY76
ID Q7SY76 PRELIMINARY; PRT; 278 AA.
AC Q7SY76;
DT 01-OCT-2003 (TREMELrel. 25, Created)
DT 01-OCT-2003 (TREMELrel. 25, Last sequence update)
DT 01-MAR-2004 (TREMELrel. 26, Last annotation update)
DE Tef-prov protein.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Whole;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.P., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Raha S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahay J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whitling M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalilus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RA "Generation and initial analysis of more than 15,000 full-length human
RA and mouse cDNA sequences.";
RA Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: Transcription factor that binds to and transactivates
CC the TSHB promoter. Binds to a minimal DNA-binding sequence 5'-
CC [TC]AG[AG]TTA[TC]AG[AG]-3'.
CC -!- SUBUNIT: Binds DNA as a homodimer or a heterodimer. Can form a
CC heterodimer with DBP.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- INDUCTION: Accumulates according to a robust circadian rhythm (By
CC similarity).
CC -!- SIMILARITY: Belongs to the bZIP family. PAR subfamily.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Whole;
 RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
 RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
 RA Richardson P.;
 RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
 RT initiative.";
 RL Dev. Dyn. 225:384-391(2002).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Whole;
 RA Klein S., Strausberg R.;
 RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: Belongs to the BZIP family.
 DR ENBL; BC054981; AAH54981.1; -;
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0003677; F:DNA binding; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro; IPR004827; TF_BZIP.
 DR SMART; SM00338; BRLZ; 1;
 DR PROSITE; PS0217; BZIP; 1.
 KW DNA-binding; Nuclear protein.
 SQ SEQUENCE 278 AA; 31536 MW; 7D475119E3F24509 CRC64;

 Query Match 44.8%; Score 13; DB 2; Length 278;
 Best Local Similarity 100.0%; Pred. No. 7e-05;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 Qy 1 RAAFLKENTALR 13
 Db 239 RAAFLKENTALR 251

 RESULT 9
 Q632P9 PRELIMINARY; PRT; 298 AA.
 AC Q632P9
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE Hypothetical protein.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
 RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
 RA Richardson P.;
 RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
 RT initiative.";
 RL Dev. Dyn. 225:384-391(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RX PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Touchman J.W., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RA Klein S., Gerhard D.S.;
 RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
 DR ENBL; BC082861; AAH82861.1; -;
 KW Hypothetical protein.
 SQ SEQUENCE 298 AA; 33649 MW; D544B58FCEA882D4 CRC64;

 Query Match 44.8%; Score 13; DB 2; Length 298;
 Best Local Similarity 100.0%; Pred. No. 7.5e-05;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 Qy 1 RAAFLKENTALR 13
 Db 253 RAAFLKENTALR 265

 RESULT 10
 Q6R212 PRELIMINARY; PRT; 188 AA.
 AC Q6R212
 DT 03-JUL-2004 (TrEMBLrel. 27, Created)
 DT 03-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE D-binding protein 3.
 GN Names=Dbp;
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Wistar; TISSUE=Brain;
 RA Klugmann M., Leitchlein C.B., Doring M.J.;
 RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: Belongs to the BZIP family.
 DR ENBL; AY518349; AAR9622.1; -;
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0003677; F:DNA binding; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro; IPR004827; TF_BZIP.
 DR SMART; SM00338; BRLZ; 1.
 DR PROSITE; PS0217; BZIP; 1.
 KW DNA-binding; Nuclear protein.
 SQ SEQUENCE 188 AA; 21040 MW; EBCED3EA67B59478 CRC64;

 Query Match 31.0%; Score 9; DB 2; Length 188;
 Best Local Similarity 100.0%; Pred. No. 0.5;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 Qy 1 RAAFLKEXEN 9
 Db 149 RAAFLKEXEN 157

 RESULT 11
 Q6R213 PRELIMINARY; PRT; 225 AA.
 ID Q6R213
 AC Q6R213;

DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE D-binding protein 2.
 GN Name=Dbp;
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Wistar; TISSUE=Brain;
 RC Klugmann M., Leitchlein C.B., During M.J.;
 RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: Belongs to the bZIP family.
 DR EMBL; AY518348; AAR99621.1; -.
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0003677; F:DNA binding; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro; IPR004827; TF_bZIP.
 DR SMART; SM00338; BRLZ; 1.
 DR PROSITE; PS50217; BZIP; 1.
 KW DNA-binding; Nuclear protein.
 SQ SEQUENCE 225 AA; 24605 MW; 53378EB8DA0B903E CRC64;
 Query Match 31.0%; Score 9; DB 2; Length 225;
 Best Local Similarity 100.0%; Pred.No. 0.59;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 RAAFLFLEKEN 9
 Db 186 RAAFLFLEKEN 194
 RESULT 12
 ID - DBP HUMAN STANDARD; PRT; 325 AA.
 AC Q10586;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE D-site-binding protein (Albumin D box-binding protein) (TAXREB302).
 GN Name=DBP;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=95137580; PubMed=7835883;
 RA Khatib Z.A., Inaba T., Valentine M., Look A.T.;
 RT "Chromosomal localization and cDNA cloning of the human DBP and TEF
 genes.";
 RL Genomics 23:344-351(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=96374825; PubMed=87861133; DOI=10.1006/geno.1996.0295;
 RA Shuttler G., Glasco T., Kang X., Korneluk R., Mueller C.R.;
 RT "Genomic structure of the human D-site binding protein (DBP) gene.";
 RL Genomics 34:334-339(1996).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RC MEDLINE=97264341; PubMed=9110174;
 RA Yu W., Andersson B., Worley K.C., Muzny D.M., Ding Y., Liu W.,
 RA Ricafrente J.Y., Wentland M.A., Lennon G., Gibbs R.A.;
 RT "Large-scale concatenation cDNA sequencing.";
 RL Genome Res. 7:353-358(1997).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Ovary;
 RC MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altshul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan F.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettman M., Madan A.C., Shevchenko Y., Bouffard G.G.,
 RA Whiting M., Touchman J.W., Green E.D., Dickson M.C.,
 RA Blakesley R.W., Grilwood J., Schmutz J., Myers R.M., Sanchez A.,
 RA Rodriguez A.C., Krzywinski M.I., Skalska U., Smalius D.E.,
 RA Butterfield Y.S.N., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [5]
 RP SEQUENCE OF 132-325 FROM N.A.
 RC MEDLINE=93246252; PubMed=8482542; DOI=10.1016/0378-1119(93)90375-D;
 RA Nyunoya H., Morita T., Sato T., Honma S., Tsujimoto A., Shimotohno K.;
 RT "Cloning of a cDNA encoding a DNA-binding protein TAXREB302 that is
 specific for the tax-responsive enhancer of HTLV-I.";
 RL Gene 126:251-255(1993).
 RN [6]
 RP REVISIONS.
 RA Nyunoya H.;
 RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
 RN [7]
 RP REVIEW.
 RC MEDLINE=99439627; PubMed=10508692; DOI=10.1016/S0959-437X(99)00009-X;
 RA Brown S.A., Schibler U.;
 RT "The ins and outs of circadian timekeeping.";
 RL Curr. Opin. Genet. Dev. 9:588-594(1999).
 CC -!- FUNCTION: This transcriptional activator recognizes and binds to
 the sequence 5'-RTTAYGTAAT-3', found in the promoter of genes such
 as albumin, CYP2A4 and CYP2A5. It is not essential for circadian
 rhythm generation, but modulates important clock output genes. May
 be a direct target for regulation by the circadian pacemaker
 component clock. May affect circadian period and sleep regulation.
 CC -!- SUBUNIT: Binds DNA as a homodimer or a heterodimer. Can form a
 heterodimer with TEF.
 CC -!- SUBCELLULAR LOCATION: Nuclear.
 CC -!- TISSUE SPECIFICITY: Ubiquitously expressed. Expressed in the
 suprachiasmatic nuclei (SCN) and in most peripheral tissues, with
 a strong circadian rhythmicity.
 CC -!- SIMILARITY: Belongs to the bZIP family. PAR subfamily.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 DR EMBL; U06936; AAA81374.1; -.
 DR EMBL; U48213; AAB18668.1; -.
 DR EMBL; U48212; AAB18668.1; JOINED.
 DR EMBL; U79283; AAB50219.1; -.
 DR EMBL; BC011965; AAH11965.1; -.
 DR EMBL; D28468; BAA05833.1; -.
 DR PIR; A55558; A55558.
 DR TRANSFAC; T04875; -.
 DR Genew; HGNC:2697; DBP.
 DR H-InvDB; HIX0015298; -.
 DR MIM; 124097; -.
 DR GO; GO:0003702; F:RNA polymerase II transcription factor acti. . .; TAS.
 DR GO; GO:0006357; P:regulation of transcription from Pol II pro. . .; TAS.
 DR InterPro; IPR004827; TF_bZIP.
 DR SMART; SM00338; BRLZ; 1.

```

DR PROSITE; PS50217; BZIP; 1.
DR PROSITE; PS0036; BZIP_BASIC; FALSE NEG.
KW Activator; Biological Rhythms; DNA-Binding; Nuclear protein;
KW Transcription regulation.
FT DOMAIN 129 135 Poly-Pro.
FT DOMAIN 188 239 Pro-rich (Proline/Acidic region (PAR)).
FT DNA_BIND 258 277 Basic motif.
FT DOMAIN 290 304 Leucine-zipper.
FT CONFLICT 179 179 S -> T (in Ref. 1).
FT CONFLICT 245 245 R -> K (in Ref. 5).
SQ SEQUENCE 325 AA; A6933CE21399ECF3 CRC64;

Query Match 31.0%; Score 9; DB 1; Length 325;
Best Local Similarity 100.0%; Pred. No. 0.81;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RAAFLEKEN 9
Db 286 RAAFLEKEN 294

RESULT 13
DBP MOUSE
ID DBP MOUSE STANDARD; PRT; 325 AA.
AC Q60925; Q8VCX3;
DT 15-JUL-1998 (Rel. 36, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DT 23-OCT-2004 (Rel. 45, Last annotation update)
DE D-size-binding protein (Albumin D box-binding protein).
GN Name=Dbp;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SV; TISSUE=Liver;
RA Lee Y.H., Oguchi H., Gonzalez F.J.;
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heish F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Tohiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
RA Boeak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Vallalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madao A., Rodrigues S., Sanchez A.,
RA Whiting M., Madao A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzyzanski M.I., Skalski U., Smalley D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [3]
RP CIRCADIEN INDUCTION.
RX PubMed=8617210;
RA Fontallaz P., Ossipow V., Wanner G., Schibler U.;
RA "The two PAR leucine zipper proteins, TEF and DBP, display similar
RT circadian and tissue-specific expression, but have different target
RT promoter preferences.";
RL EMO J. 15:351-362 (1996).
RN [4]
RP IMPLICATION IN EPILEPSY.

```

```

RX PubMed=15175240; DOI=10.1101/gad.301404;
RA Gachon F., Fontallaz P., Damiola F., Gos P., Kodama T., Zakany J.,
RA Duboule D., Petit B., Tafti M., Schibler U.;
RT "The loss of circadian PAR bZIP transcription factors results in
RT epilepsy.";
RL Genes Dev. 18:1397-1412 (2004).
CC -I- FUNCTION: This transcriptional activator recognizes and binds to
CC the sequence 5'-RTAVGTAAY-3', found in the promoter of genes such
CC as albumin, CYP2A4 and CYP2A5. It is not essential for circadian
CC rhythm generation, but modulates important clock output genes. May
CC be a direct target for regulation by the circadian pacemaker
CC component clock. May affect circadian period and sleep regulation
CC (By similarity).
CC -I- SUBUNIT: Binds DNA as a homodimer or a heterodimer. Can form a
CC heterodimer with TEF (By similarity).
CC -I- SUBCELLULAR LOCATION: Nuclear.
CC -I- TISSUE SPECIFICITY: Expressed in the suprachiasmatic nuclei (SCN)
CC and in most peripheral tissues, with a strong circadian
CC rhythmicity.
CC -I- INDUCTION: Accumulates according to a robust circadian rhythm in
CC liver and kidney. In liver nuclei, the amplitude of daily
CC oscillation has been estimated to be >50-fold, and 2-fold in the
CC brain.
CC -I- MISCELLANEOUS: Mice deficient for all three PAR bZIP proteins
CC (DBP, HLF and TEF) display a dramatically shortened life span and
CC are highly susceptible to generalized spontaneous and audiogenic
CC epilepsies (due for example to the noise of a vacuum cleaner) that
CC are frequently lethal. The down-regulation of pyridoxal kinase
CC (pdxk) expression in these mice may participate in this seizure
CC phenotype.
CC -I- SIMILARITY: Belongs to the bZIP family. PAR subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U29762; AAA73924.1; -.
CC EMBL; BC018323; AAH18323.1; -.
CC GMD; MGI:94866; Dbp.
CC InterPro; IPR004827; TF bZIP.
CC SMART; SM00338; BRU; 1.
CC PROSITE; PS50217; BZIP; 1.
CC PROSITE; PS00036; BZIP_BASIC; FALSE NEG.
KW Activator; Biological rhythms; DNA-Binding; Nuclear protein;
FT DOMAIN 129 135 Poly-Pro.
FT DOMAIN 156 161 Poly-Ser.
FT DOMAIN 188 239 Pro-rich (Proline/Acidic region (PAR)).
FT DNA_BIND 258 277 Basic motif.
FT DOMAIN 290 304 Leucine-zipper.
FT CONFLICT 60 60 S -> T (in Ref. 1).
FT CONFLICT 109 109 L -> F (in Ref. 1).
SQ SEQUENCE 325 AA; 34379 MW; B2B2A3E091845A16 CRC64;

Query Match 31.0%; Score 9; DB 1; Length 325;
Best Local Similarity 100.0%; Pred. No. 0.81;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RAAFLEKEN 9
Db 286 RAAFLEKEN 294

RESULT 14
DBP RAT
ID DBP RAT STANDARD; PRT; 325 AA.
AC P16443;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)

```

25-OCT-2004 (Rel. 45, Last annotation update)
D-site-binding protein (Albumin D box-binding protein) (D site albumin
promoter binding protein 1).

Name=Dbp;
Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
NCBI_TaxID=10116;
[1]

SEQUENCE FROM N.A.

SPRAIN=Wistar; TISSUE=Liver;
RX MEDLINE=90235277; PubMed=2331750;

RA Mueller C.R., Maire P., Schibler U.;

RT "DBP, a liver-enriched transcriptional activator, is expressed late in
ontogeny and its tissue specificity is determined

posttranscriptionally.";

RL Cell 61:279-291(1990).

RN [2]

REVISTONS.

RA Mueller C.R., Maire P., Schibler U.;

RL Cell 65:915-915(1991).

RN [3]

RP REVIEW.
RX MEDLINE=99439627; PubMed=10508692; DOI=10.1016/S0959-437X(99)00009-X;

RA Brown S.A., Schibler U.;

RL "The ins and outs of circadian timekeeping.";

RT Curr. Opin. Genet. Dev. 9:588-594(1999).

CC -1- FUNCTION: This transcriptional activator recognizes and binds to
the sequence 5'-ATTAYGTAAV-3' found in the promoter of genes such

as albumin, CYP2A4 and CYP2A5. It is not essential for circadian
rhythm generation, but modulates important clock output genes. May

be a direct target for regulation by the circadian pacemaker

component clock. May affect circadian period and sleep regulation

(By similarity).

CC -1- SUBUNIT: Binds DNA as a homodimer or a heterodimer. Can form a

heterodimer with TEF.

CC -1- SUBCELLULAR LOCATION: Nuclear.

CC -1- TISSUE SPECIFICITY: Expressed in the suprachiasmatic nuclei (SCN)
and in most peripheral tissues, with a strong circadian

rhythmicity.

CC -1- DEVELOPMENTAL STAGE: Expressed late in ontogeny.

CC -1- SIMILARITY: Belongs to the bZIP family. PAR subfamily.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; J03179; AA41083.1; -.

PIR; A34894; A34894.

TRANSFAC; T00183; -.

RGD; 2491; Dbp.

DR InterPro; IPR004827; TF bZIP.

DR SMART; SM00338; BRLZ; 1.

DR PROSITE; PS00217; bZIP; 1.

DR PROSITE; PS00036; bZIP_BASIC; FALSE NEG.

KW Activator; Biological rhythms; DNA-binding; Nuclear protein;

Transcription regulation.

FT DOMAIN 129 135 Poly-Pro.

FT DOMAIN 156 161 Poly-Ser.

FT DOMAIN 188 239 Pro-rich (Proline/Acidic region (PAR)).

FT DNA BIND 258 277 Basic motif.

FT DOMAIN 290 304 Leucine-zipper.

SEQUENCE 325 AA; 34436 MW; D9B2A53FF18455B2 CRC64;

Query Match 31.0%; Score 9; DB 1; Length 325;

Best Local Similarity 100.0%; Pred. No. 0.81;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY

1 RAAFLKEN 9

Db 286 RAAFLKEN 294
|||||

RESULT 15

RIB7_METTH

ID _RIB7_METTH

AC 026337; STANDARD; PRT; 216 AA.

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 05-JUL-2004 (Rel. 44, Last annotation update)

DE Putative 5-amino-6-(5-phosphoribosylamino)uracil reductase

(EC 1.1.1.193) (HTP reductase).

DE Ordered locus names=WH235;

OS Methanobacterium thermoautotrophicum.

OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;

OC Methanobacteriaceae; Methanothermobacter.

OC NCBI_TaxID=187420;

RN [1]

SEQUENCE FROM N.A.

RC STRAIN=Delta H;

RX MEDLINE=98037514; PubMed=9371463;

RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,

Alredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,

Harrison D., Hoang L., Keagle P., Lamm W., Pothier B., Qiu D.,

Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,

Jiwani N., Caruso A., Bush D., Safer H., Patwell D., Prabakar S.,

McDougall S., Shimer G., Goyal A., Pietrovski S., Church G.M.,

Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;

"Complete genome sequence of Methanobacterium thermoautotrophicum

deltaH: functional analysis and comparative genomics.";

RL J. Bacteriol. 179:7135-7155(1997).

CC -1- CATALYTIC ACTIVITY: 5-amino-6-(5-phosphoribosylamino)uracil +

NADPH(+) = 5-amino-6-(5-phosphoribosylamino)uracil + NADPH.

CC -1- PATHWAY: Riboflavin biosynthesis; third step.

CC -1- SIMILARITY: Belongs to the HTP reductase family.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; AE000810; AAB84741.1; -.

PIR; C69129; C69129.

DR InterPro; IPR006401; Rib reduct_arch.

DR InterPro; IPR002734; RibD_C.

DR Pfam; PF01872; RibD_C; 1.

DR TIGRFAMs; TIGR01508; rib_reduct_arch; 1.

DR TIGRFAMs; TIGR00227; ribD_Cterm; 1.

KW Complete proteome; Hypothetical protein; NADP; Oxidoreductase;

Riboflavin biosynthesis

SEQUENCE 216 AA; 23576 MW; 436561C089DF41A6 CRC64;

Query Match 27.6%; Score 8; DB 1; Length 216;

Best Local Similarity 100.0%; Pred. No. 5.7;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 15 RVAELRKR 22

|||||

Db 108 RVAELRKR 115

Search completed: July 26, 2005, 12:19:37

Job time : 67.6528 secs

This page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 27, 2005, 10:08:29 ; Search time 1201.08 Seconds
(without alignments)
1169.945 Million cell updates/sec

Title: US-10-624-218-2

Perfect score, 29
Sequence: 1 RAAPLEXENTALTRVAELRRVGRCRNI 29

Scoring table:
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Word size: 1

Total number of hits satisfying chosen parameters: 9406972

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:
-MODE=frame+2n.model -DEV=xlp
-Q=/cgn2_1/USPO_spool_p/US10624218/runat_26072005_121435_3685/app.query.fasta_1.796
-DB=GenEmbl -QMT=fastcap -SUFFIX=olip2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=oligo -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc
-NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10624218 @CGN 1 1 6148 @runat_26072005_121435_3685 -NCPU=6 -ICPU=3
-NO_MMAP -LARGOQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*

1: gb.ba.*
2: gb.hcg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sv.*
13: gb.un.*
14: gb.vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	14	48.3	446	AX305813	Sequence
2	14	48.3	665	10 PSTEF	Y11149 P.sungorus
3	14	48.3	817	10 SS8745	SS8745 thyrotroph
4	14	48.3	855	9 HSU06935	U06935 Human thyro

5	14	48.3	909	9	CR541827	CR541827 Homo sapi
6	14	48.3	974	9	CR456592	CR456592 Homo sapi
7	14	48.3	1000	9	HSU44059	U44059 Human thyro
8	14	48.3	1003	6	CQ730263	CQ730263 Sequence
9	14	48.3	1153	5	CGU09221	U09221 Gallus gall
10	14	48.3	1304	10	AF194420	AF194420 Mus muscu
11	14	48.3	2823	9	AK091916	AK091916 Homo sapi
12	14	48.3	3103	10	AY540632	AY540632 Mus muscu
13	14	48.3	3808	9	HSM805945	BS337848 Homo sapi
14	14	48.3	4018	10	BC017689	BC017689 Mus muscu
15	14	48.3	4064	10	AY540631	AY540631 Mus muscu
16	14	48.3	4222	10	BC036982	BC036982 Mus muscu
17	14	48.3	4375	9	BC039258	BC039258 Homo sapi
18	14	48.3	4382	9	BC042476	BC042476 Homo sapi
19	14	48.3	61834	2	AC068105	AC068105 Homo sapi
20	14	48.3	102151	9	HS979N1	AL035659 Human DNA
21	14	48.3	151983	2	AC027178	AC027178 Homo sapi
22	14	48.3	219714	2	AC096601	AC096601 Rattus no
23	14	48.3	262820	2	AC102103	AC102103 Mus muscu
24	13	44.8	2017	5	BC082861	BC082861 Xenopus l
25	13	44.8	2594	5	BC042248	BC042248 Xenopus l
26	13	44.8	2654	5	BC054981	BC054981 Xenopus l
27	9	31.0	807	10	AY518349	AY518349 Rattus no
28	9	31.0	865	10	BC084676	BC084676 Mus muscu
29	9	31.0	918	10	AY518348	AY518348 Rattus no
30	9	31.0	978	9	BT006836	BT006836 Homo sapi
31	9	31.0	978	12	BT007806	BT007806 Synthetic
32	9	31.0	1370	6	CQ714530	CQ714530 Sequence
33	9	31.0	1385	6	CQ861501	CQ861501 Sequence
34	9	31.0	1385	9	HUMDSITE2	U48213 Human D-sit
35	9	31.0	1391	9	HUMDBPD	D28468 Homo sapien
36	9	31.0	1403	9	HSU06936	U06936 Human album
37	9	31.0	1476	10	BC018323	BC018323 Mus muscu
38	9	31.0	1478	9	BC011965	BC011965 Homo sapi
39	9	31.0	1480	9	HSU79283	U79283 Human album
40	9	31.0	1671	6	AX827439	AX827439 Sequence
41	9	31.0	1671	10	RATDBPA	J03179 Rat D-bindi
42	9	31.0	1749	10	BC064094	BC064094 Mus muscu
43	9	31.0	5527	10	MMU29762	U29762 Mus musculu
44	9	31.0	38149	9	AC022154	AC022154 Homo sapi
45	9	31.0	61797	2	AC023094	AC023094 Homo sapi

ALIGNMENTS

RESULT 1
AX305813
LOCUS AX305813 Sequence 564 from Patent WO0188188.
DEFINITION AX305813
ACCESSION AX305813
VERSION AX305813.1 GI:17645203
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE 1
AUTHORS Ishikawa,K., Asai,S., Takahashi,Y., Nagata,T. and Ishii,Y.
TITLE Method for examining ischemic conditions
JOURNAL Patent: WO 0188188-A 564 22-NOV-2001;
School Juridical Person Nihon University (JP)
FEATURES
source
1..446
/organism="Mus musculus"
/mol_type="unassigned DNA"
/db_xref="taxon:10090"

ORIGIN

Alignment Scores:
Pred. No.: 7.33e-05 Length: 446
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0

Query Match:	48.28%	Indels:	0
DB:	6	Gaps:	0
US-10-624-218-2 (1-29) x AX305813 (1-446)			
Qy	1 ArgAlaAlaPheLeuGluLysGluAsnThrAlaLeuArgThr 14		
Db	88 CGGCGAGGTTCTCTGGAGAGGAGACACAGCCCTGCGGACG 129		
RESULT 2			
PSTEF			
LOCUS	PSTEF	665 bp	mRNA linear ROD 12-FEB-1997
DEFINITION	P.sungorus mRNA for thyrotroph embryonic factor, partial.		
ACCESSION	Y11149		
VERSION	Y11149.1 GI:1841478		
KEYWORDS	TEF; thyrotroph embryonic factor.		
SOURCE	Phodopus sungorus (Djungarian hamster)		
ORGANISM	Phodopus sungorus		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae; Phodopus.		
REFERENCE	1		
AUTHORS	Bockmann, J.		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 665)		
AUTHORS	Bockmann, J.		
TITLE	Direct Submission		
JOURNAL	Submitted (10-FEB-1997) J. Bockmann, Institute of Anatomy, University of Muenster, Vesaliusweg 2-4, 48149 Muenster, FRG		
FEATURES	Location/Qualifiers		
source	1..665		
	/organism="Phodopus sungorus"		
	/mol_type="mRNA"		
	/db_xref="taxon:10044"		
CDS	<1..>665		
	/note="RT-PCR fragment"		
	/codon_start=3		
	/product="thyrotroph embryonic factor (TEF)"		
	/protein_id="CAA72036.1"		
	/db_xref="GI:1841479"		
	/db_xref="GOA:P97516"		
	/db_xref="UniProt/Swiss-Prot:P97516"		
	/translation="LEEDSAAASTMAVSASLMPPIWDKTIPTYDGESFHLFYMDLDF LLENGIPASPHLAQNLLPVALEGGESASSSTASPSSTAVFQSPSTVSSSTESL EKERTSPIDPNCVEVDVNFNPPADLVSSVPGELNPRKHFAEEDLKQPMIK KAKKVFVDEQDEKYTRKKNVNAKRSDARLKENQITIRAAFLKENTALRTE VAD"		
ORIGIN			
Alignment Scores:			
Pred. No.:	0.000102	Length:	665
Score:	14.00	Matches:	14
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	48.28%	Indels:	0
DB:	10	Gaps:	0
US-10-624-218-2 (1-29) x PSTEF (1-665)			
Qy	1 ArgAlaAlaPheLeuGluLysGluAsnThrAlaLeuArgThr 14		
Db	612 CGGCGAGCCTTCTCTGGAGAGGAGACACAGCCCTGCGGACG 653		
RESULT 3			
S58745			
LOCUS	S58745	817 bp	mRNA linear ROD 07-MAY-1993
DEFINITION	thyrotroph embryonic factor=leucine zipper transcription factor [rats, pituitary, mRNA, 817 nt].		
ACCESSION	S58745		
VERSION	S58745.1 GI:237084		
KEYWORDS	Rattus sp.		
SOURCE	Rattus sp.		
ORGANISM	Rattus sp.		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.		
REFERENCE	1 (bases 1 to 817)		
AUTHORS	Drolet,D.W., Scully,K.M., Simmons,D.M., Wegner,M., Chu,K.T., Swanson,L.W. and Rosenfeld,M.G.		
TITLE	TEF, a transcription factor expressed specifically in the anterior pituitary during embryogenesis, defines a new class of leucine zipper proteins		
JOURNAL	Genes Dev. 5 (10), 1739-1753 (1991)		
MEDLINE	92009166		
LOCUS	1916262		
REMARK	GenBank staff at the National Library of Medicine created this entry [NCBI gibbsq 58745] from the original journal article.		
FEATURES	Location/Qualifiers		
source	1..817		
	/organism="Rattus sp."		
	/mol_type="mRNA"		
	/db_xref="taxon:10118"		
gene	1..817		
	/gene="thyrotroph embryonic factor, TEF"		
CDS	32..817		
	/gene="thyrotroph embryonic factor, TEF"		
	/note="leucine zipper transcription factor; TEF"		
	/codon_start=1		
	/product="thyrotroph embryonic factor"		
	/protein_id="AAB20032.1"		
	/db_xref="GI:237085"		
	/translation="MENPPRETRLDKEKLEKLEDESAAASTMAVSASLMPPIWDKTIPTYDGESFHLFYMDLDFLENGIPASPHLAQNLLPVALEGGESASSSTASPSSTAVFQSPSTVSSSTESLSTAIQFQSPSTVSSSEKEREKTPIDPNCVEVDVNFNPPADLVSSVPGGELNPRKHFAEEDLKQPMIKKAKKVFVDEQDEKYTRKKNVNAKRSDARLKENQITIRAAFLKENTALRTEVAELRKEVGKCTIVSKYETKYGL"		
ORIGIN			
Alignment Scores:			
Pred. No.:	0.000122	Length:	817
Score:	14.00	Matches:	14
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	48.28%	Indels:	0
DB:	10	Gaps:	0
US-10-624-218-2 (1-29) x S58745 (1-817)			
Qy	1 ArgAlaAlaPheLeuGluLysGluAsnThrAlaLeuArgThr 14		
Db	695 CGGCGAGCATTTCTGGAAGAAGACACAGCCCTGCGGACG 736		
RESULT 4			
HSU06935			
LOCUS	HSU06935	855 bp	mRNA linear PRI 16-NOV-1995
DEFINITION	Human thyrotroph embryonic factor (TEF) mRNA, complete cds.		
ACCESSION	U06935		
VERSION	U06935.1 GI:606796		
KEYWORDS	Homo sapiens (human)		
SOURCE	Homo sapiens		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1 (bases 1 to 855)		
AUTHORS	Khatib,Z.A., Inaba,T., Valentine,M. and Look,A.T.		
TITLE	Chromosomal localization and cDNA cloning of the human DBP and TEF genes		
JOURNAL	Genomics 23 (2), 344-351 (1994)		
MEDLINE	95137580		
PUBMED	7835883		
REFERENCE	2 (bases 1 to 855)		
AUTHORS	Inaba,T.		
TITLE	Direct Submission		
JOURNAL	Submitted (18-FEB-1994) Toshiya Inaba, St. Jude Children's Research Hospital, 332 North Laderdale, Memphis, TN 38105-0318, USA		
FEATURES	Location/Qualifiers		

```
source
1..855
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cell_line="UOC-B1"
/cell_type="B-cell precursor"
1..855
/gene="TEF"
<1..50
/gene="TEF"
51..836
/gene="TEF"
/function="transcription factor"
/codon_start=1
/evidence=experimental
/product="thyrotroph embryonic factor"
/protein_id="AA81373.1"
/db_xref="GI:606797"
/translation="MENPREARLDEEKGKLEDEAAASTMAVSASLMPPIWDKT
IPYDGESFHELYMDLEFLENGIPASPHLAHLLLPVALEKESASSSTAPSS
STAIPOQSETVSSSELEKETSPIDPNCVEVDVNFNPDADLVSSVFGGSELFN
PRKHFAEDLKPQMIKKAKVFPDQKDEKYWTRKKNNVAAKRSRDARLKENQ
ITIRAAFLKENTALRTEVALRKEVGKCKTIVSKYKGYPL"
837..>855
/gene="TEF"

3'UTR
FEATURES
            source
            Location/Qualifiers
1..909
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="RZPD0834D0532D"
/clone_lib="Human Full ORF Clones Gateway(TM) - RZPD"
/lab_host="DH5Alpha"
/notes="Vector: pDONR201, Site_1: attP1; Site_2: attP2"
1..909
/gene="TEF"
1..>909
/gene="TEF"
/codon_start=1
/protein_id="CAG46626.1"
/db_xref="GI:49456611"
/translation="MSDAGGKKPPVDPOAGPGPGRAAGRGLSGSPPLVLKILME
NPPREARLDKSGKEKLEDEAAASTMAVSASLMPPIWDKTI PYDGESFHELYMDLD
EFLLENGIPASPHLAHLLLPVALEKESASSSTAPSSSTAIPOQSETVSSSTES
SLEKRETPSPIDPNCVEVDVNFNPDADLVSSVFGGSELFNPRKHFAEDLKPQPM
IKKAKVFPDQKDEKYWTRKKNNVAAKRSRDARLKENQITIRAAFLKENTALR
TEVALRKEVGKCKTIVSKYKGYPL"

ORIGIN
Alignment Scores:
Pred. No.: 0.000133 Length: 909
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 48.28% Indels: 0
DB: 9 Gaps: 0

US-10-624-218-2 (1-29) x CR541827 (1-909)
QY 1 ATGAlaAlaPheLeuGluLysGluAsnThrAlaLeuArgThr 14
|||||
Db 714 CGGCAGCCTTCTCGAGAGGAGAGAACACAGCCCTGCGGACG 755

RESULT 5
CR541827 909 bp mRNA linear PRI 29-JUN-2004
LOCUS Full ORF shuttle clone, Gateway(TM), complete cds.
DEFINITION Homo sapiens full open reading frame cDNA clone RZPD0834D0532D for
             gene TEF, thyrotrophic embryonic factor; complete cds, without
             stopcodon.
ACCESSION CR541827
VERSION 1 GI:49456610
KEYWORDS Full ORF shuttle clone, Gateway(TM), complete cds.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 909)
AUTHORS Halleck,A., Ebert,L., Mkoondinya,M., Schick,M., Eisenstein,S.,
          Neubert,P., Katrang,K., Schatten,R., Shen,B., Henze,S., Mar,W.,
          Korn,B., Zuo,D., Hu,Y. and LaBaer,J.
          Cloning of human full open reading frames in Gateway(TM) system
          entry vector (pDONR201)
          Unpublished
          2 (bases 1 to 909)
          Halleck,A., Ebert,L., Mkoondinya,M., Schick,M., Eisenstein,S.,
          Neubert,P., Katrang,K., Schatten,R., Shen,B., Henze,S., Mar,W.,
          Korn,B., Zuo,D., Hu,Y. and LaBaer,J.
          Direct Submission
          Submitted (28-JUN-2004) RZPD Deutsches Ressourcenzentrum fuer
          Genomforschung GmbH, Im Neuenheimer Feld 580, D-69120 Heidelberg,
          Germany
          RZPD; RZPD0834D0532D, ORFNo 3756
          www.rzpd.de/cgi-bin/products/c1.cgi?CloneID=RZPD0834D0532D RZPDLIB;
          Human Full ORF Clones Gateway(TM) - RZPD (kan-resist.) RZPD LIB No.
          834

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
RZPD; RZPD0834D0532D, ORFNo 3756
www.rzpd.de/cgi-bin/products/c1.cgi?CloneID=RZPD0834D0532D RZPDLIB;
Human Full ORF Clones Gateway(TM) - RZPD (kan-resist.) RZPD LIB No.
834

www.rzpd.de/cgi-bin/products/showLib.pl.cgi/response?libNo=834
www.rzpd.de/products/orfclones/
Contact: Inge Arian
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH,
Heubnerweg 6, D-14059 Berlin, Germany
Tel: +49 30 32639 100
Fax: +49 30 32639 111
www.rzpd.de
This clone is available from RZPD;
Contact RZPD (customer.service@rzpd.de) for further information.
Clone name at Harvard Institute of Proteomics
(www.hip.harvard.edu): FLH130899.01L
This CDS clone is part of a collection of human full ORF clones
jointly established and verified by the Harvard Institute of
Proteomics (HIP) and RZPD.
This CDS has been cloned without stopcodon.
The CDS has been inserted into pDONR201 via a BP Clonase(TM)
reaction. Additional sequence has been added in front of the start
codon: att..AAAAA GCA GGC TCC ACC (ATG).
The last codon is followed by the 3' att site: GACCCAGCTTCTT. .att
The clone is validated by full sequence check.
Compared to the reference sequence NM_003216 (GI:34486096) we did
not find any amino acid exchanges.
Clone distribution: http://www.rzpd.de/products/orfclones/.
```

JOURNAL	Submitted (29-DEC-1995) Stephen P. Hunger, Pediatrics, University of Colorado Health Sciences Center, UCHSC, Box C229, 4200 E 9th Ave, Denver, CO 80262, USA		
	Location/Qualifiers		
FEATURES	source		
	1..1000		
gene	/organism="Homo sapiens"		
	/mol_type="mRNA"		
CDS	/db_xref="taxon:9606"		
	/chromosome="22"		
FEATURES	/map="22q13.2"		
	/cell_line="ALL cell line HB11.19"		
gene	1..1000		
	/gene="TEF"		
CDS	37..948		
	/gene="TEF"		
FEATURES	/note="TEF; similar to product encoded by Gallus gallus vbp gene, GenBank Accession number U09221; PAR-BZIP transcription factor"		
	/codon_start=1		
ORIGIN	/product="thyrotroph embryonic factor"		
	/protein_id="AAB06497.1"		
Alignment Scores:	/db_xref="GI:1399386"		
	/translation="MSDAGGKKPPVDPQAGPGPGRAAGRLSGSFFLVKLME		
Pred. No.:	NPPEARLDKEKGEKLEDEAAASTWVASASLMPPIWDKTI PYDGSFHLFYMDD		
	EFLLENGIPASPTLHALLLLPVALECKESASSTASPPSSSTAI FQPSSTVSTES		
Score:	SLEKERETSPIDPNCVEVDVNFDPADLVLSVPGGELFNPRKHKEAEDLKQPM		
	IKKAKKVFVPDEQDEKYWRKKNNVAAKSRDARLKENQITIRAAFLFLEKENTAIR		
Percent Similarity:	TEVAELRKEVGKCKTIVSKYETKYGL"		
	Length: 1000		
Best Local Similarity:	Matches: 14		
	Conservative: 0		
Query Match:	Mismatches: 0		
	Indels: 0		
DB:	Gaps: 0		
	US-10-624-218-2 (1-29) x HSU44059 (1-1000)		
Qy	1 ArgAlaAlaPheLeuGluLysGluAsnThrAlaLeuArgThr 14		
Db	826 CGGGCAGCCTTCTCGGAGGAGGAGACACAGCCCTGGGACG 867		
RESULT 8	CQ730263		
	LOCUS		
DEFINITION	Sequence 16197 from Patent WO02068579.		
	CQ730263		
ACCESSION	CQ730263.1		
	GI:42303641		
VERSION	CQ730263.1		
	Homo sapiens (human)		
KEYWORDS	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
SOURCE	Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.		
	1		
REFERENCE	Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.		
	Kits, such as nucleic acid arrays, comprising a majority of		
AUTHORS	humanexons or transcripts, for detecting expression and other uses		
	thereof		
JOURNAL	Patent: WO 02068579-A 16197 06-SEP-2002;		
	PE Corporation (NY) (US)		
FEATURES	Location/Qualifiers		
	1..1003		
ORIGIN	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
Alignment Scores:	/db_xref="taxon:9606"		
	Length: 1003		
Pred. No.:	0.000145		
	Matches: 14		
Score:	Conservative: 0		
	Percent Similarity: 100.00%		
Query Match:	100.00%		
	Indels: 0		
DB:	Gaps: 0		
	US-10-624-218-2 (1-29) x CR456592 (1-974)		
Qy	1 ArgAlaAlaPheLeuGluLysGluAsnThrAlaLeuArgThr 14		
Db	803 CGGGCAGCCTTCTCGGAGGAGGAGACACAGCCCTGGGACG 844		
RESULT 7	HSU44059		
	LOCUS		
DEFINITION	Human thyrotroph embryonic factor (TEF) mRNA, complete cds.		
	U44059		
ACCESSION	U44059.1		
	GI:1399385		
VERSION	U44059.1		
	Homo sapiens (human)		
KEYWORDS	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
SOURCE	Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.		
	1		
REFERENCE	Hunger, S.P., Li, S., Fall, M.Z., Naumovski, L. and Cleary, M.L.		
	The proto-oncogene HLF and the related basic leucine zipper protein		
AUTHORS	TEF display highly similar DNA-binding and transcriptional		
	regulatory properties		
JOURNAL	Blood 87 (11), 4607-4617 (1996)		
	MEDLINE		
PUBMED	96219638		
	8639829		
REFERENCE	2 (bases 1 to 1000)		
	Hunger, S.P.		
AUTHORS	Direct Submission		
	TITLE		

Db 868 CGGCAGCCTTCCTTGAGAAAGAGATACGGCCCTGAGACG 909
|||||
RESULT 10
AF194420 1304 bp mRNA linear ROD 19-MAY-2000
LOCUS Mus musculus thyrotroph embryonic factor (Tef) mRNA, complete cds.
DEFINITION AF194420
ACCESSION AF194420
VERSION AF194420.1 GI:7939635
KEYWORDS
SOURCE
ORGANISM Mus musculus (house mouse)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1304)
AUTHORS Krueger, D.A., Warner, E.A. and Dowd, D.R.
TITLE Involvement of thyrotroph embryonic factor in calcium-mediated
regulation of gene expression
JOURNAL J. Biol. Chem. 275 (19), 14524-14531 (2000)
MEDLINE 20261582
PUBMED 10799536
REFERENCE 2 (bases 1 to 1304)
AUTHORS Dowd, D.R., Krueger, D.A. and Warner, B.A.
TITLE Direct Submision
JOURNAL Submitted (12-OCT-1999) Pharmacology, Case Western Reserve
University School of Medicine, 10900 Euclid Ave., Cleveland, OH
44106-4965, USA
FEATURES
Location/Qualifiers
1..1304
/organism="Mus musculus"
/mol_type="mRNA"
/strain="BALB/c"
/db_xref="taxon:10090"
/sex="male"
/tissue_type="brain"
/dev_stage="adult"
1..1304
/gene="Tef"
/note="TEF; PAR-bZIP transcription factor"
/codon_start=1
/product="thyrotroph embryonic factor"
/protein_id="AAF0830.1"
/db_xref="GI:7939636"
/translation="MSDAGGGKKPVEPQAGPGRAAGLGGSPFLVLKXLMENP
PRETRLDKKGKLEDESAASAMVSAASLMPPIWDKTI PYDGESFHELYMDLDEF
LLENGIPASPTHLAQNLLPVAELGKESASSTASPPSSSTAI FQPSSTVSTSSLL
EKERTSPIDPSCVEVDVNFNPDADLVLSVPGGELFNPKHFAEDLKPQMIK
KAKKVFVPEQDEKQYTRKKNNVAAKRRDARKLKENQITIRAAFLKENTALRTE
VAELRKEVGKCKTIVSKYETKYGPL"
ORIGIN
Alignment Scores:
Pred. No.: 0.00018 Length: 1304
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 48.28% Indels: 0
DB: 10 Gaps: 0
US-10-624-218-2 (1-29) x AF194420 (1-1304)
Qy 1 ArgAlaAaPheLeuGluLysGluAsnThrAlaLeuArgThr 14
|||||
Db 874 CGGCAGCCTTCCTTGAGAAAGAGATACGGCCCTGAGACG 915
|||||
RESULT 11
AK091916 2823 bp mRNA linear PRI 30-JAN-2004
LOCUS Homo sapiens cDNA FLJ34597 fis, clone KIDNE2009367, highly similar
DEFINITION to THYROTROPH EMBRYONIC FACTOR.
ACCESSION AK091916

Best Local Similarity: 100.00% Mismatches: 0
Query Match: 48.28% Indels: 0
DB: 6 Gaps: 0
US-10-624-218-2 (1-29) x CQ730263 (1-1003)
Qy 1 ArgAlaAaPheLeuGluLysGluAsnThrAlaLeuArgThr 14
|||||
Db 829 CGGCAGCCTTCCTTGAGAAAGAGATACGGCCCTGAGACG 870
|||||
RESULT 9
GGU09221 1153 bp mRNA linear VRT 30-NOV-1995
LOCUS Gallus gallus White Leghorn beta/beta isoform of vitellogenin
DEFINITION binding protein (vbp) mRNA, complete cds.
ACCESSION U09221
VERSION U09221.1 GI:483937
KEYWORDS
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
REFERENCE 1 (bases 1 to 1153)
AUTHORS Burch, J.B. and Davis, D.L.
TITLE Alternative promoter usage and splicing options result in the
differential expression of mRNAs encoding four isoforms of chicken
VBP, a member of the PAR subfamily of bZIP transcription factors
JOURNAL Nucleic Acids Res. 22 (22), 4733-4741 (1994)
MEDLINE 95075656
PUBMED 7984425
REFERENCE 2 (bases 1 to 1153)
AUTHORS Burch, J.B.
TITLE Direct Submision
JOURNAL Submitted (26-APR-1994) John B.E. Burch, Fox Chase Cancer Center,
7701 Burholme Ave., Philadelphia, PA 19111, USA
FEATURES
Location/Qualifiers
1..1153
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="White Leghorn"
/db_xref="taxon:9031"
/cell_type="fibroblast"
/clone_lib="Vennstrom"
/dev_stage="embryonic day 10"
1..1153
/gene="vbp"
/note="vbp"
/codon_start=1
/product="vitellogenin binding protein (VBP), beta/beta
isoform"
/protein_id="AAA82156.1"
/db_xref="GI:483938"
/translation="MSVCNAGSGAALDPPEVLKSLLEYSLPWTMTDKKKIKLEE
DEAAASTAVASLMPPIWDKTI PYDGESFHELYMDLDEFLENGIPSPHLDLQ
NPLMPVAKLEKEPASATGSPVSSSTAVYQSEASSTSPONERTPSPIPDPC
KVWFNFNPDADLVLSVPGGELFNPKHKTEDLKPQMIKAKKVFVPEQDE
KYWTRKKNNVAAKRRDARKLKENQITIRAAFLKENTALRTEVAELRKEVGRCNI
VSKYTRYGPFPLSDSE"
ORIGIN
Alignment Scores:
Pred. No.: 0.000163 Length: 1153
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 48.28% Indels: 0
DB: 5 Gaps: 0
US-10-624-218-2 (1-29) x GGU09221 (1-1153)
Qy 1 ArgAlaAaPheLeuGluLysGluAsnThrAlaLeuArgThr 14
|||||

VERSION	AK091916.1 GI-21750394									
KEYWORDS	oligo capping; f1s (full insert sequence).									
SOURCE	Homo sapiens									
ORGANISM	Homo sapiens									
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.									
AUTHORS	1 Ota,T., Suzuki,Y., Nishikawa,T., Otsuki,T., Sugiyama,T., Irie,R., Wakamatsu,A., Hayaashi,K., Sato,H., Nagai,K., Kimura,K., Makita,H., Sekine,M., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T., Ishii,S., Yamamoto,J., Saito,K., Kawai,Y., Isono,Y., Nakamura,Y., Nagahari,K., Murakami,K., Yasuda,T., Iwayanagi,T., Wagatsuma,M., Shiratori,A., Sudo,H., Hosoiri,T., Kaku,Y., Kodaira,H., Kondo,H., Sugawara,M., Takahashi,M., Kanda,K., Yokoi,T., Furuya,T., Kikkawa,E., Omura,K., Abe,K., Kamiyama,K., Katsuta,N., Sato,K., Tanikawa,M., Yanazaki,M., Ninomiya,K., Ishibashi,T., Yamashita,H., Hiraoka,S., Chiba,Y., Ishida,S., Ono,Y., Takiguchi,S., Watanabe,S., Yosida,M., Hotta,T., Kusano,J., Kanehori,K., Takahashi-Fujii,A., Hara,H., Tanase,T., Nomura,Y., Togiya,S., Komai,F., Hara,R., Takeuchi,K., Arita,M., Imose,N., Musashino,K., Yuuki,H., Oshima,A., Sasaki,N., Aotsuka,S., Yoshikawa,Y., Matsunawa,H., Ichihara,T., Shiohata,N., Sano,S., Moriya,S., Momiyama,H., Satoh,N., Takami,S., Terashima,Y., Suzuki,O., Nakagawa,S., Senoh,A., Mizoguchi,H., Goto,Y., Shimizu,F., Wakebe,H., Hishigaki,H., Watanabe,T., Sugiyama,A., Takemoto,M., Kawakami,B., Yamazaki,M., Watanabe,K., Kumagai,A., Itakura,S., Fukuzumi,Y., Fujimori,Y., Fujimori,Y., Komiyama,M., Tashiro,H., Tanigami,A., Fujizawa,T., Ono,F., Yamada,K., Fujii,Y., Ozaki,K., Hirao,M., Ohmori,Y., Kawabata,A., Hikiji,T., Kobatake,N., Inagaki,H., Ikema,Y., Okamoto,S., Okitani,R., Kawakami,T., Noguchi,S., Itoh,T., Shigetani,K., Senba,T., Matsumura,K., Nakajima,Y., Miuno,T., Morinaga,M., Sasaki,M., Togashi,T., Oyama,M., Hata,H., Watanabe,M., Komatsu,T., Mizushima-Sugano,J., Satoh,T., Shirai,Y., Takahashi,Y., Nakagawa,K., Okumura,K., Nagase,T., Nomura,N., Kikuchi,H., Masuho,Y., Yamashita,R., Nakai,K., Yada,T., Nakamura,Y., Ohara,O., Isogai,T. and Sugano,S.									
TITLE	Complete sequencing and characterization of 21,243 full-length human cDNAs									
JOURNAL	Nat. Genet. 36 (1), 40-45 (2004)									
PUBMED	14702039									
REFERENCE	2 Tashiro,H., Yamazaki,M., Watanabe,K., Kumagai,A., Itakura,S., Fukuzumi,Y., Fujimori,Y., Komiyama,M., Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H., Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K., Wagatsuma,M., Murakami,K., Kanehori,K., Takahashi-Fujii,A., Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y., Sugano,S., Nagahari,K., Masuho,Y., Nagai,K. and Isogai,T.									
TITLE	NEDO human cDNA sequencing project									
JOURNAL	Unpublished									
REFERENCE	3 (bases 1 to 2823)									
AUTHORS	Isogai,T. and Yamamoto,J.									
TITLE	Direct Submission									
JOURNAL	Submitted (04-JUL-2002) Takao Isogai, FUJ Project (HRI Team) ; 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)									
COMMENT	NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.									
FEATURES	Location/Qualifiers									
source	1. .2823									
	/organism="Homo sapiens"									
	/mol_type="mRNA"									
	/db_xref="taxon:9606"									
	/clone="KIDNE2009367"									
	/issue_type="kidney"									
	/clone_lib="KIDNE2"									
ORIGIN	Alignment Scores:									
	Pred. No.: 0.000373 Length: 3103									
	Score: 14.00 Matches: 14									
	Percent Similarity: 100.00% Conservativeness: 0									
	Best Local Similarity: 100.00% Mismatches: 0									
	Query Match: 48.28% Indels: 0									
	DB: 9 Gaps: 0									
ORIGIN	US-10-624-218-2 (1-29) x AK091916 (1-2823)									
QY	1 ArgAlaAlaPheLeuGluIysGluAsnThrAlaLeuArgThr 14									
Db	778 CGGGCAGCCTTCTCTGGAGAGGAGAACACAGCCCTCGCGACG 819									
RESULT 12	AY540632 3103 bp mRNA linear ROD 12-APR-2004									
LOCUS	AY540632 Mus musculus thymotroph embryonic factor beta isoform (Tef) mRNA, complete cds, alternatively spliced.									
DEFINITION	AY540632.1 GI:42768795									
ACCESSION	AY540632									
VERSION	AY540632.1									
KEYWORDS	Mus musculus (house mouse)									
SOURCE	Mus musculus									
ORGANISM	Mus musculus									
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.									
AUTHORS	1 (bases 1 to 3103)									
TITLE	Zhou,J., Hoggatt,A.M. and Herring,B.P.									
JOURNAL	Activation of the smooth muscle-specific telokin gene by thymotroph embryonic factor (TEF)									
PUBMED	J. Biol. Chem. 279 (16), 15929-15937 (2004)									
REFERENCE	2 (bases 1 to 3103)									
AUTHORS	Zhou,J., Hoggatt,A.M. and Herring,B.P.									
TITLE	Direct Submission									
JOURNAL	Submitted (30-JAN-2004) Cellular and Integrative Physiology, Indiana University School of Medicine, 635 Barnhill Drive, MS2067, Indianapolis, IN 46202, USA									
FEATURES	Location/Qualifiers									
source	1. 3103									
	/organism="Mus musculus"									
	/mol_type="mRNA"									
	/db_xref="taxon:10090"									
	/chromosome="15"									
	/tissue_lib="bladder"									
gene	1. 3103									
	/gene="Tef"									
CDS	6. .863									
	/gene="Tef"									
	/note="transcription factor; alternatively spliced"									
	/codon_start=1									
	/product="thymotroph embryonic factor beta isoform"									
	/protein_id="AA045600.1"									
	/cd_xref="GI:42768796"									
	/translation="MSSCSIQGVAPMDPEVLKSLLEHSLPWSEKKADKEKKEKLE EDSSAATWVASLMPPIWDKTIIPYDGESFLETMDLDEFLENGIPASPTHLAON LLAPVALEKGSASSSTAIQFQSETSVTSSESLERKERTIPSPIDPSCVE VDVNFNPDADLVLSVPGGELFNPRFAEDLKQPQMKKAKKVFVPEQKDEKY WTKRKNNVAAKSRDARLKENQITIRAAFLKENTALTALTEVALRKEVKGCKTIYS KYETKYGLP"									
ORIGIN	Alignment Scores:									
	Pred. No.: 0.000373 Length: 3103									
	Score: 14.00 Matches: 14									
	Percent Similarity: 100.00% Conservativeness: 0									
	Best Local Similarity: 100.00% Mismatches: 0									
	Query Match: 48.28% Indels: 0									
	DB: 10 Gaps: 0									

/note="cloning vector: pME18SFL3"

US-10-624-218-2 (1-29) x AY540632 (1-3103)

Qy 1 ArgAlaAlaPheLeuGluLysGluAsnThrAlaLeuArgThr 14
|||||

Db 741 CGGCAGCGCTTCTGGAGAGAGACACACGCCCTGGGACG 782

RESULT 13
HSM805945 3808 bp mRNA linear PRI 17-JUN-2003
LOCUS Homo sapiens mRNA; cDNA DKFZp686D1282 (from clone DKFZp686D1282).
ACCESSION BX537848
KEYWORDS BX537848.1 GI:31873831
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 3808)
Wambutt, R., Heubner, D., Mewes, H.W., Weil, B., Amid, C., Osanger, A.,
Fobo, G., Han, M. and Wiemann, S.
Direct Submission
Submitted (17-JUN-2003) MIPS, Ingolstaedter Landstr.1, D-85764
Neuherberg, GERMANY
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by AGOWA (Berlin/Germany) within the cDNA sequencing
consortium of the German Genome Project.
This clone (DKFZp686D1282) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
information about the clone and the sequencing project is available
at http://mips.gsf.de/proj/cDNA/.

FEATURES
Source

1..3808
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/map="22q13.2"
/clone="DKFZp686D1282"
/tissue_type="human retina"
/clone_lib="686 (synonym: hicc3). Vector pSport1_Sfi; host
DH108; sites SfiIa + SfiIB"
/dev_stage="adult"
1..3808
/gene="DKFZp686D1282"
1..425
/gene="DKFZp686D1282"
/note="Thyrotroph embryonic factor, N-terminus truncated"
/codon_start=3
/product="hypothetical protein"
/protein_id="CAD97856.1"
/db_xref="GI:31873832"
/translation="KERTPSPIDPNCVEVDVNFDPADLVLSVPGELPNRKHK
FAEDLPQPMIKKAKKVFVDEQDKIWTNRKNVAAKRRDRARLKENQITIRA
AFLEKENTALRTVAELRKEVGKCKTIIVSKYETKYGL"
3770..3775
/gene="DKFZp686D1282"
3787
/gene="DKFZp686D1282"

ORIGIN

Alignment Scores:
Pred. No.: 0.000443 Length: 3808
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 48.28% Indels: 0
DB: Gaps: 0

US-10-624-218-2 (1-29) x HSM805945 (1-3808)

Qy 1 ArgAlaAlaPheLeuGluLysGluAsnThrAlaLeuArgThr 14
|||||

303 CGGCAGCGCTTCTGGAGAGAGACACACGCCCTGGGACG 344

RESULT 14
BC017689

LOCUS
DEFINITION

Mus musculus thyrotroph embryonic factor, transcript variant 2,
mRNA (cDNA clone MGC:19233 IMAGE:4242534), complete cds.
4018 bp mRNA linear ROD 30-JUN-2004

ACCESSION BC017689.1 GI:17389268

VERSION
KEYWORDS

SOURCE
MGC.

ORGANISM
Mus musculus (house mouse)

REFERENCE
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 4018)

Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L.H., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heide, F.,
Diatchenko, L., Narusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Uedin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalobon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettman, M., Madan, A., Young, A.C., Shvachenko, Y.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Green, E.D.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalusz, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932

2 (bases 1 to 4018)

Strausberg, R.

Direct Submission

TITLE

Submitted (03-DEC-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

NIH-MGC Project URL: http://mgc.nci.nih.gov

Contact: MGC help desk

Email: cgapbs-remail.nih.gov

Tissue Procurement: Jeffrey E. Green, M.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305

Web site: http://www.shgc.stanford.edu

Contact: (Dickson, Mark) mcd@paxil.stanford.edu

Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov

Series: IRAC Plate: 25 Row: j Column: 24

This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 23821037.

Location/Qualifiers

1..4018

/organism="Mus musculus"

/mol_type="mRNA"

/strain="FVB/N"

/db_xref="taxon:10090"

/clone="MGC:19233 IMAGE:4242534"

/tissue_type="Kidney, normal. 5 month old male mouse."

/clone_lib="NCI CGAP_Kid14"

/lab_host="DH108"

/note="Vector: pCMV-SPORT6"

FEATURES

Source

```
gene
1. .4018
/genes="Tef"
/db_xref="LocusID:21685"
/db_xref="MGI:98663"
4. .825
/genes="Tef"
/codon_start=1
/product="thyrotroph embryonic factor, isoform 2"
/protein_id="AAH17689.1"
/db_xref="GI:17389269"
/db_xref="LocusID:21685"
/db_xref="MGI:98663"
/translation="MDMPVLSLLEHSLPWSKKADKKGKLEDESAASTMAV
SASLMPPIWKTIPYDGESPHLYMDLDEFLLENGIPASPTHLAQNLILPVALEQKE
SASSTASPPSSSTALFPQSETVSSSTESSELEKERTSPIDPSCVEVDVNVNPDADL
VLSVPGGELFNPRKHFABEDLKQPQMIKKAKVFPDQKDEKWTWRKKNNVAAK
RSRDARLKENQITIRAAFLKENTALTVEALRKEVGKCKTIVSKYETKYGPL"

ORIGIN
Alignment Scores:
Pred. No.: 0.000463 Length: 4018
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 48.28% Indels: 0
DB: 10 Gaps: 0

US-10-624-218-2 (1-29) x BC017689 (1-4018)
QY 1 ArgAlaAlaPheLeuGlulysGluAsnThrAlaLeuArgThr 14
Db 703 CGGGCAGCCTTCCTGGAGAGGAGACACAGCCCTGCGGACG 744

RESULT 15
AY540631 4064 bp mRNA linear ROD 12-APR-2004
LOCUS Mus musculus thyrotroph embryonic factor alpha isoform (Tef) mRNA,
DEFINITION complete cds, alternatively spliced.
ACCESSION AY540631
VERSION AY540631.1 GI:42768793
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 4064)
Zhou, J., Hoggatt, A.M. and Herring, B.P.
Activation of the smooth muscle-specific telokin gene by thyrotroph
embryonic factor (TEF)
J. Biol. Chem. 279 (16), 15929-15937 (2004)
14702338
JOURNAL
PUBMED
2 (bases 1 to 4064)
Zhou, J., Hoggatt, A.M. and Herring, B.P.
Direct Submission
AUTHORS
TITLE Submitted (30-JAN-2004) Cellular and Integrative Physiology,
JOURNAL Indiana University School of Medicine, 635 Barnhill Drive, MS2067,
Indianapolis, IN 46202, USA
INDIANAPOLIS, IN 46202, USA
FEATURES
source
1. .4064
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/chromosome="15"
/tissue_lib="bladder"
1. .4064
/genes="Tef"
143. .1048
/genes="Tef"
/notes="transcription factor; alternatively spliced"
/codon_start=1
/product="thyrotroph embryonic factor alpha isoform"
/protein_id="AA545599.1"
/db_xref="GI:42768794"

/translation="MSDAGGGKPPVPQAGPGRAAGRGSLSGSPFLVLKLMENP
PRETRLDKKEKLEEDESAASTMAVSAASLMPPIWDKTIPIDGESPHLYMDLDEF
LLENGIPASPTHLAQNLILPVALEQKESASSTASPPSSSTALFPQSETVSSSTES
EKERTSPIDPSCVEVDVNVNPDADLVLSSVPGGELFNPRKHFABEDLKQPQMIK
KAKVFPDQKDEKWTWRKKNNVAAKRSRDARLKENQITIRAAFLKENTALTVEAL
RKEVGKCKTIVSKYETKYGPL"

ORIGIN
Alignment Scores:
Pred. No.: 0.000468 Length: 4064
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 48.28% Indels: 0
DB: 10 Gaps: 0

US-10-624-218-2 (1-29) x AY540631 (1-4064)
QY 1 ArgAlaAlaPheLeuGlulysGluAsnThrAlaLeuArgThr 14
Db 926 CGGGCAGCCTTCCTGGAGAGGAGACACAGCCCTGCGGACG 967

Search completed: July 27, 2005, 13:54:06
Job time : 1203.08 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 27, 2005, 10:05:45 ; Search time 232.403 Seconds
(without alignments)
738.685 Million cell updates/sec

Title: US-10-624-218-2

Perfect score: 29

Sequence: 1 RAAFLKENTALTTRVAELRKVRGRCRN 29

Scoring table:

OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Word size: 1

Total number of hits satisfying chosen parameters: 8774627

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters: -DB=frame+ p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO spool_p/US10624218/runat_26072005_121435_3677/app.query.fasta_1.796
-DB=N Geneseq_16Dec04 -QMT=fastcap -SUFFIX=olip2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=oligo -TRANS=human40.cdd
-LIST=45 -DOCALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFWT=pt0 -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10624218 @CGN 1.1 1241 @runat_26072005_121435_3677 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : N Geneseq_16Dec04: *
1: geneseqn1980s: *
2: geneseqn1990s: *
3: geneseqn2000s: *
4: geneseqn2001as: *
5: geneseqn2001bs: *
6: geneseqn2002as: *
7: geneseqn2002bs: *
8: geneseqn2003as: *
9: geneseqn2003bs: *
10: geneseqn2003cs: *
11: geneseqn2003ds: *
12: geneseqn2004as: *
13: geneseqn2004bs: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	14	48.3	446	6	Abi99561 Mouse isc
2	14	48.3	999	13	Adq87162 Human tum
3	14	48.3	999	13	Adq87502 Human tum
4	14	48.3	1000	12	Adi82448 Human mod
5	14	48.3	1000	13	Acn40954 Tumour-as

	6	14	48.3	4480	5	ABA16225	AbA16225 Human ner
	7	14	48.3	5240	5	ABA16224	AbA16224 Human ner
C	8	9	31.0	217	12	ACH91032	ACH91032 Human gen
	9	9	31.0	466	9	ACH35960	ACH35960 Human end
	10	9	31.0	480	13	ADQ49232	AdQ49232 Novel can
	11	9	31.0	491	12	ACH88921	ACH88921 Human gen
	12	9	31.0	546	12	ACH75221	ACH75221 Human gen
C	13	9	31.0	546	12	ACH77332	ACH77332 Human gen
	14	9	31.0	1385	12	ADP13309	AdP13309 Renal cel
	15	9	31.0	1385	13	ADR52783	AdR52783 Drug ther
	16	9	31.0	1403	12	ADI82446	AdI82446 Human mod
	17	9	31.0	1403	12	ADO19733	AdO19733 Human PRO
	18	9	31.0	1403	13	ADP54858	AdP54858 Human PRO
	19	9	31.0	1671	12	ADP72558	AdP72558 Renal tox
	20	9	31.0	2678	12	ADQ87010	AdQ87010 Human tum
	21	9	31.0	2678	12	ADQ83706	AdQ83706 Human tum
	22	9	31.0	2678	12	ADQ85851	AdQ85851 Human tum
	23	9	31.0	5801	5	ABA18280	AbA18280 Human ner
C	24	8	27.6	57	12	ADL90017	AdL90017 Gluconoba
C	25	8	27.6	1365	5	AAH67383	AaH67383 C glutami
	26	8	27.6	1517	4	AAF71641	Aaf71641 Corynebac
	27	8	27.6	1893	12	ADL90008	AdL90008 Gluconoba
	28	8	27.6	2025	5	AAH67382	AaH67382 C Glutami
	29	8	27.6	2155	10	ADD13430	Add13430 C. glutam
	30	8	27.6	11769	6	ABS53606	AbS53606 Human Mut
C	31	8	27.6	28580	10	ADD45921	Add45921 Rat gene
	32	8	27.6	58215	11	ACN44100	Acn44100 Mouse gen
C	33	8	27.6	96594	9	ADA02756	Ada02756 Human MOR
C	34	8	27.6	96594	10	ADB72494	Adb72494 Human MOR
C	35	8	27.6	96594	10	ADC85236	Adc85236 Human Mor
C	36	8	27.6	96594	12	ADM74351	Adm74351 Human car
C	37	8	27.6	101505	11	ACN44694	Acn44694 Human gen
	38	8	27.6	215974	12	ADQ97523	Adq97523 Human can
	39	8	27.6	349980	5	AAH68531	Aah68531 C glutami
	40	7	24.1	197	3	AAC17847	Aac17847 Human sec
	41	7	24.1	264	10	ADH82372	Adh82372 Enterococ
C	42	7	24.1	290	6	ABV95929	Abv95929 Human pan
	43	7	24.1	315	12	ADL87447	Adl87447 DNA up-re
	44	7	24.1	315	12	ADL87446	Adl87446 DNA up-re
	45	7	24.1	326	6	ABQ89421	Abq89421 Human pro

ALIGNMENTS

RESULT 1
ABI99561
ID ABI99561 standard; CDNA; 446 BP.

AC ABI99561;

DT 07-MAR-2002 (first entry)

DE Mouse ischaemic condition related cDNA sequence SEQ ID NO:564.

XX Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;

KW vasospastic ischaemia; ischaemic condition; ischaemic disease; ss.

OS Mus musculus.

PN WO200188188-A2.

XX 22-NOV-2001.

XX 18-MAY-2001; 2001WO-JP004192.

XX 18-MAY-2000; 2000JP-00145977.

XX (UYNI-) UNIV NIHON SCHOOL JURIDICAL PERSON.

XX Ishikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y;

XX WPI; 2002-034733/04.

XX

Db	826	CGGCGAGCCTTCTCTGGAGAGGAGAACACAGCCCTCGGACG	867
RESULT 5			
ACN40954			
ID	ACN40954	standard; cDNA; 1000 BP.	
XX	AC		
XX	AC	ACN40954;	
XX	18-NOV-2004	(first entry)	
XX	Tumour-associated	antigenic target (TAT) CDNA DNA326982, SEQ ID NO:6077.	
XX	Tumour-associated	antigenic target; TAT; human; overexpression; cancer;	
XX	tumour; diagnosis;	cell proliferative disorder; breast cancer;	
XX	colorectal cancer;	lung cancer; ovarian cancer; liver cancer;	
XX	central nervous system	cancer; bladder cancer; pancreatic cancer;	
XX	cervical cancer;	melanoma; leukaemia; hybridisation probe;	
XX	chromosome identification;	chromosome mapping; gene mapping;	
XX	gene therapy; cytostatic;	gene; ss.	
XX	Homo sapiens.		
XX	WO2004030615-A2.		
XX	15-APR-2004.		
XX	29-SEP-2003;	2003WO-US028547.	
XX	02-OCT-2002;	2002US-0414971P.	
XX	(GETH)	GENENTECH INC.	
XX	Wu TD, Zhang Z, Zhou Y;		
XX	WPI; 2004-347921/32.		
XX	P-PSDB; ABW82367.		
XX	New tumor-associated	antigenic target polypeptides and nucleic acids,	
XX	useful in preparing a	medicament for treating or detecting a	
XX	proliferative disorder,	e.g. breast, lung, colorectal, ovarian or	
XX	prostate cancer or tumor.		
XX	Claim 1; SEQ ID NO	6077; 7273pp; English.	
XX	The invention relates	to human tumour-associated antigenic target (TAT)	
XX	polypeptides, and their	related nucleic acids. The TAT polypeptides are	
XX	overexpressed in cancer	tissues compared to normal tissues, and may thus	
XX	serve as effective targets	for the diagnosis and treatment of cancer in	
XX	mammals. The invention	also relates to nucleic acid and polypeptide	
XX	sequences at least 80%	identical to the TAT nucleic acids and	
XX	polypeptides; expression	vectors and host cells comprising a TAT nucleic	
XX	acid; an antibody specific	for a TAT polypeptide; a peptide or organic	
XX	molecule which binds to	a TAT polypeptide; fusion proteins comprising a	
XX	TAT polypeptide; and	methods and compositions for the treatment or	
XX	diagnosis of cancer in	mammals. TAT polypeptides, nucleic acids,	
XX	antibodies, antagonists,	binding molecules and compositions are useful	
XX	for diagnosing or treating	a cell proliferative disorder associated with	
XX	increased TAT expression,	particularly cancers such as breast cancer,	
XX	colorectal cancer, lung	cancer, ovarian cancer, liver cancer, bladder	
XX	cancer, pancreatic cancer,	cervical cancer, cancers of the central	
XX	nervous system, melanoma	and leukaemia. TAT nucleic acids may further be	
XX	used as hybridisation	probes, in chromosome and gene mapping, in	
XX	chromosome identification	and in gene therapy. The present sequence	
XX	represents a TAT nucleic	acid of the invention	
XX	Sequence 1000 BP;	231 A; 312 C; 307 G; 150 T; 0 U; 0 Other;	
XX	Alignment Scores:		
XX	Pred. No.:	5,79e-05	Length: 1000
XX	Score:	14.00	Matches: 14
XX	Percent Similarity:	100.00%	Conservative: 0
XX	Best Local Similarity:	100.00%	Mismatches: 0

Query Match:	48.28%	Indels:	0
DB:	13	Gaps:	0
US-10-624-218-2	(1-29) x ACN40954	(1-1000)	
QY	1	ArgAlaAlaPheLeuGluLysGluAsnThrAlaLeuArgThr	14
DB	826	CGGCGAGCCTTCTCTGGAGAGGAGAACACAGCCCTCGGACG	867
RESULT 6			
ABAI6225			
ID	ABAI6225	standard; DNA; 4480 BP.	
XX	AC		
XX	AC	ABAI6225;	
XX	23-JAN-2002	(first entry)	
XX	Human nervous system	related polynucleotide SEQ ID NO 8556.	
XX	Human; nootropic;	neuroprotective; cytostatic; dermatological; virucide;	
XX	immunosuppressive;	antiinflammatory; anti-HIV; antibacterial; vulnerary;	
XX	antiparkinsonian;	antisickling; antianaemic; antiarthritic; cancer;	
XX	antirheumatic;	hepatotropic; cerebroprotective; antiinflammatory;	
XX	antiallergic;	antidiabetic; antiulcer; anticonvulsant; antifungal;	
XX	antiparasitic;	cardiant; immune disorder; cardiovascular disorder;	
XX	neurological disease;	infection; nephrotropic; gene therapy; vaccine; ds.	
XX	Homo sapiens.		
XX	WO200159063-A2.		
XX	16-AUG-2001.		
XX	17-JAN-2001;	2001WO-US001334.	
XX	31-JAN-2000;	2000US-0179065P.	
XX	04-FEB-2000;	2000US-0180828P.	
XX	24-FEB-2000;	2000US-0184664P.	
XX	02-MAR-2000;	2000US-0186350P.	
XX	16-MAR-2000;	2000US-0189874P.	
XX	17-MAR-2000;	2000US-0190076P.	
XX	18-APR-2000;	2000US-0198123P.	
XX	19-MAY-2000;	2000US-0205515P.	
XX	07-JUN-2000;	2000US-0209467P.	
XX	28-JUN-2000;	2000US-0214886P.	
XX	30-JUN-2000;	2000US-0215135P.	
XX	07-JUL-2000;	2000US-0216647P.	
XX	07-JUL-2000;	2000US-0216880P.	
XX	11-JUL-2000;	2000US-0217487P.	
XX	11-JUL-2000;	2000US-0217496P.	
XX	14-JUL-2000;	2000US-0218290P.	
XX	26-JUL-2000;	2000US-0220963P.	
XX	26-JUL-2000;	2000US-0220964P.	
XX	14-AUG-2000;	2000US-0224518P.	
XX	14-AUG-2000;	2000US-0224519P.	
XX	14-AUG-2000;	2000US-0225213P.	
XX	14-AUG-2000;	2000US-0225214P.	
XX	14-AUG-2000;	2000US-0225266P.	
XX	14-AUG-2000;	2000US-0225267P.	
XX	14-AUG-2000;	2000US-0225268P.	
XX	14-AUG-2000;	2000US-0225270P.	
XX	14-AUG-2000;	2000US-0225447P.	
XX	14-AUG-2000;	2000US-0225757P.	
XX	14-AUG-2000;	2000US-0225758P.	
XX	14-AUG-2000;	2000US-0225759P.	
XX	18-AUG-2000;	2000US-0226279P.	
XX	22-AUG-2000;	2000US-0226681P.	
XX	22-AUG-2000;	2000US-0226868P.	
XX	23-AUG-2000;	2000US-0227182P.	
XX	23-AUG-2000;	2000US-0227009P.	
XX	30-AUG-2000;	2000US-0228944P.	
XX	01-SEP-2000;	2000US-0229287P.	
XX	01-SEP-2000;	2000US-0229343P.	

PR 01-SEP-2000; 2000US-02293344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234223P.
PR 25-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235835P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 13-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 20-OCT-2000; 2000US-0242221P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.

PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250391P.
PR 01-DEC-2000; 2000US-0251160P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-541565/60.
XX
XX Nucleic acids encoding 3224 human nervous system antigen polypeptides,
PT useful for preventing, diagnosing and/or treating nervous system cancers
PT and metastases.
XX
PS Disclosure; SEQ ID NO 8556; 1701pp + Sequence Listing; English.
XX
XX The invention relates to novel genes (ABA11004-ABA21534) and proteins
CC (AB114678-AB118001) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in
CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
CC cardiovascular disorders such as myocardial ischaemia; (d) wound healing
CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
CC infectious diseases such as viral, bacterial, fungal and parasitic
CC infections. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 4480 BP; 913 A; 1195 C; 1214 G; 1158 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 0.000231 Length: 4480
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 48.28% Indels: 0
DB: 5 Gaps: 0

US-10-624-218-2 (1-29) x ABA16225 (1-4480)

QY 1 ArgAlaAlaPheLeuGluLysGluAsnThrAlaLeuArgThr 14
|||
Db 1745 CGGGCAGCCTTCTCTGGAGAGGAGACACAGCCCTCGGACG 1786

RESULT 7

✓

PR 17-NOV-2000; 2000US-0249299P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 01-DEC-2000; 2000US-0251160P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 05-DEC-2000; 2000US-0256719P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 08-DEC-2000; 2000US-0251990P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA Rosen CA, Barash SC, Ruben SM;
 FI WPI; 2001-541565/60.
 XX Nucleic acids encoding 3224 human nervous system antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating nervous system cancers
 PT and metastases.
 XX Disclosure; SEQ ID NO 8555; 1701pp + Sequence Listing; English.
 PS The invention relates to novel genes (AB11004-ABA21534) and proteins
 XX (AB14678-AB18001) useful for preventing, treating or ameliorating
 CC medical conditions e.g. by protein or gene therapy. The genes are
 CC isolated from a range of human tissues disclosed in the specification.
 CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in
 CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
 CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
 CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
 CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
 CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
 CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
 CC cardiovascular disorders such as myocardial ischaemias; (d) wound healing
 CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
 CC infectious diseases such as viral, bacterial, fungal and parasitic
 CC infections. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 5240 BP; 1079 A; 1354 C; 1452 G; 1355 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 0.000267 Length: 5240
 Score: 14.00 Matches: 14
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 48.28% Indels: 0
 DB: 5 Gaps: 0

US-10-624-218-2 (1-29) x ABA16224 (1-5240)
 QY 1 ArgAlaAlaPheLeuGluLysGluSerThrAlaLeuArgThr 14
 DB 1745 CGGCGAGCCTTCCTGGAAGGAGACACAGCCCTCGGACG 1786

RESULT 8
 ACH91032/c
 ID ACH91032 standard; DNA; 217 BP.
 XX
 AC ACH91032;
 XX
 DT 29-JUL-2004 (first entry)
 XX
 DE Human genome derived single exon probe #24227.
 XX
 KW Human; probe; ss; gene expression; single exon probe; microarray;

KW alternative splicing event; genomic alteration.
 XX Homo sapiens.
 XX US2003194704-A1.
 XX 16-OCT-2003.
 XX 03-APR-2002; 2002US-00029386.
 XX 03-APR-2002; 2002US-00029386.
 XX (PENN/) PENN S G.
 PA (RANK/) RANK D R.
 PA (HANZ/) HANZEL D K.
 XX Penn SG, Rank DR, Hanzel DK;
 XX WPI; 2004-119264/12.
 XX New human genome-derived single exon nucleic acid probes useful for human
 PT gene expression analysis, for identifying or characterizing alternative
 PT splicing events, for assessing genomic alterations or as tools for
 PT surveying tissues.
 XX Claim 1; SEQ ID NO 24227; 80pp; English.
 XX The invention relates to a nucleic acid probe for measuring human gene
 CC expression, comprising any of the 27,400 fully defined nucleotide
 CC sequences in the specification, or their complements or fragments, and
 CC encoding at least 8 amino acids of any of the 8988 amino acid sequences
 CC fully defined in the specification. The probe is a single exon probe that
 CC hybridises under high stringency conditions to a nucleic acid molecule
 CC expressed in human cells or tissues. Also included are a spatially-
 CC addressable set of single exon nucleic acid probes for measuring human
 CC gene expression (comprising a plurality of single exon nucleic acid
 CC probes cited above, where each of the plurality of probes is separately
 CC and addressably isolatable or amplifiable from the plurality), a single
 CC exon microarray for measuring human gene expression, a method of
 CC measuring human gene expression, a vector comprising the single exon
 CC probe cited above, an ORF-encoded peptide comprising at least 8
 CC contiguous amino acids of any of the above-mentioned amino acid
 CC sequences (optionally with conservative amino acid substitutions), an
 CC isolated antibody that binds specifically to a peptide cited above,
 CC methods of selling and/or licensing single exon probes or microarrays to
 CC a customer desiring to measure gene expression, a method of providing
 CC human gene expression data by subsequence, and a computer-readable
 CC storage medium which contains a database having a plurality of records
 CC (each record including data on the expression of a single exon probe
 CC cited above. The probe, methods and apparatus are useful in gene
 CC expression analysis. The probes may be used as tools for surveying
 CC tissues to detect the presence of expressed messages that contain their
 CC specific exon, or in constructing genome-derived single exon microarrays.
 CC In addition, the probes are used in identifying and characterising
 CC alternative splicing events, in detecting and characterising gross
 CC alterations in the genomic locus that includes their exon, in assessing
 CC smaller genomic alterations, in priming the synthesis of nucleic acids,
 CC or in expressing the ORF-encoded peptide. The present sequence is a human
 CC single exon probe of the invention. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from USPIO at
 CC seqdata.uspto.gov/sequence.html?DocID=20030194704
 XX
 SQ Sequence' 217 BP; 28 A; 77 C; 66 G; 46 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 1.73 Length: 217
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 31.03% Indels: 0
 DB: 12 Gaps: 0

```
US-10-624-218-2 (1-29) x ACH91032 (1-217)
QY      1 ArgAlaAlaPheLeuGluLysGluAsn 9
      |||||
Db      123 CGGCGCGCCTTCTGGAGAGGAGAC 97

RESULT 9
ACH35960
ID ACH35960 standard; cDNA; 466 BP.
XX
AC ACH35960;
XX
DT 13-OCT-2003 (first entry)
XX
DE Human endothelial cell cDNA #4093.
XX
KW Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
KW genome mapping; biodiversity; genetic disorder.
XX
OS Homo sapiens.
XX
PN US2003073623-A1.
XX
PD 17-APR-2003.
XX
PF 30-JUL-2001; 2001US-00918995.
XX
PR 30-JUL-2001; 2001US-00918995.
XX
PA (DRMA/) DRMANAC R T.
PA (LABA/) LABAT I.
PA (STAC/) STACHE-CRAIN B.
PA (DICK/) DICKSON M C.
PA (JONE/) JONES L W.
XX
PI Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;
XX
DR WPI; 2003-615964/58.
XX
PT New polynucleotide sequences obtained from various cDNA libraries, useful
PT as hybridization probes, as oligomers for PCR, for chromosome and gene
PT mapping, in the recombinant production of protein, or in generating
PT antisense DNA or RNA.
XX
PS Claim 1; SEQ ID NO 23172; 44pp; English.
XX
CC The invention relates to an isolated polynucleotide comprising any one of
CC 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
CC determined by the technique of SBH (sequencing by hybridisation). Also
CC included is a purified polypeptide comprising a sequence corresponding to
CC a reading frame of the novel polynucleotide. The nucleic acid sequences
CC are useful in diagnostics as expressed sequence tags (EST) for
CC identifying expressed genes or for physical mapping of the human genome,
CC in forensics, in assessing biodiversity, or in identifying mutations
CC responsible for genetic disorders and other traits. The nucleotide
CC sequences are also useful as hybridisation probes, as oligomers for PCR,
CC for chromosome and gene mapping, in the recombinant production of
CC protein, or in generating antibodies specific for it. The present sequence
CC is useful for generating antibodies specific for it. The present sequence
CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20030073623
XX
SQ Sequence 466 BP; 114 A; 133 C; 129 G; 90 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 3.5 Length: 466
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 31.03% Indels: 0
DB: 9 Gaps: 0
```

```
US-10-624-218-2 (1-29) x ACH35960 (1-466)
QY      1 ArgAlaAlaPheLeuGluLysGluAsn 9
      |||||
Db      407 CGGCGCGCCTTCTGGAGAGGAGAC 433

RESULT 10
ADQ49232
ID ADQ49232 standard; DNA; 480 BP.
XX
AC ADQ49232;
XX
DT 21-OCT-2004 (first entry)
XX
DE Novel canine microarray-related DNA sequence SeqID534.
XX
KW canine microarray; drug screening; toxicity assay;
KW environmental pollutant; cellular response; gene expression profile;
KW toxic response; liver necrosis; fatty liver disease;
KW protein adduct formation; hepatitis; dog; ds.
XX
OS Canis familiaris.
XX
PN WO2004063324-A2.
XX
PD 29-JUL-2004.
XX
PF 05-MAY-2003; 2003WO-US013853.
XX
PR 03-MAY-2002; 2002US-0377240P.
XX
PA (GENE-) GENE LOGIC INC.
PA (PFIZ ) PFIZER PROD INC.
XX
PI Diggans JC, Porter M, Wei T;
XX
DR WPI; 2004-561890/54.
XX
PT New isolated nucleic acid molecule, useful for drug screening and
PT toxicity assays or for assessing the impact, including toxicity, of a
PT compound, pharmaceutical agent or environmental pollutant on a cell or
PT living organism.
XX
PS Claim 1; SEQ ID NO 534; 41pp; English.
XX
CC This invention is related to a novel isolated canine nucleic acid
CC sequences and the construction of canine microarrays containing a
CC significant portion of the canine genome. The isolated canine nucleic
CC acid sequences of the invention may be useful for drug screening and
CC toxicity assays. The invention is therefore useful for assessing the
CC impact, including toxicity, of a compound, pharmaceutical agent or
CC environmental pollutant on a cell or living organism. The methods are
CC useful for detecting genes that are up- or down-regulated in canines in a
CC disease state. The sequences are useful as diagnostic agents or markers
CC to detect a cellular response in a sample individually or as part of a
CC gene expression profile. It is also useful as a target for agents that
CC modulate gene expression or activity. The database is useful for
CC producing electronic Northern blots that allow the user to determine the cell
CC type or tissue in which a given gene is expressed and to allow
CC determination of the abundance or expression level of a given gene in a
CC particular tissue or cell. The methods are useful for determining the
CC similarity of a toxic response to one or more individual compounds. The
CC methods are useful for predicting at least one toxic response or the
CC likelihood that a compound or test agent will induce various specific
CC pathologies such as those of the liver (liver necrosis, fatty liver
CC disease, protein adduct formation or hepatitis), those of the kidney,
CC heart, brain or testes, or other pathologies associated with at least one
CC of the toxins. The methods are also useful for predicting or elucidating
CC the potential cellular pathways influenced, induced or modulated by the
CC compound or test agent due to the similarity of the expression profile
CC compared to the profile induced by a known toxin. The present sequence is
CC that of a canine DNA sequence which was claimed for use during the
```


CC production of a canine microarray of the invention.
XX
SQ Sequence 480 BP; 100 A; 151 C; 146 G; 77 T; 0 U; 6 Other;
Alignment Scores:
Pred. No.: 3.6 Length: 480
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 31.03% Indels: 0
DB: 13 Gaps: 0
US-10-624-218-2 (1-29) x ADQ49232 (1-480)
Qy 1 ArgAlaAlaPheLeuGluLysGluAsn 9
Db 142 CGGGGGGCTTCCTGGAGAGAGAGAAC 168
RESULT 11
ACH88921
ID ACH88921 standard; DNA; 491 BP.
XX
AC ACH88921;
XX
DT 29-JUL-2004 (first entry)
XX
DE Human genome derived single exon probe #22116.
XX
KW Human; probe; ss; gene expression; single exon probe; microarray;
KW alternative splicing event; genomic alteration.
XX
OS Homo sapiens.
XX
PN US2003194704-A1.
XX
PD 16-OCT-2003.
XX
PF 03-APR-2002; 2002US-00029386.
XX
PR 03-APR-2002; 2002US-00029386.
XX
PA (PENN/) PENN S G.
PA (RANK/) RANK D R.
PA (HANZ/) HANZEL D K.
XX
PI Penn SG, Rank DR, Hanzel DK;
XX
DR WPI; 2004-119264/12.
XX
PT New human genome-derived single exon nucleic acid probes useful for human
PT gene expression analysis, for identifying or characterizing alternative
PT splicing events, for assessing genomic alterations or as tools for
PT surveying tissues.
XX
PS Claim 1; SEQ ID NO 22116; 80pp; English.
XX
CC The invention relates to a nucleic acid probe for measuring human gene
CC expression, comprising any of the 27,400 fully defined nucleotide
CC sequences in the specification, or their complements or fragments, and
CC encoding at least 8 amino acids of any of the 6888 amino acid sequences
CC fully defined in the specification. The probe is a single exon probe that
CC hybridizes under high stringency conditions to a nucleic acid molecule
CC expressed in human cells or tissues. Also included are a spatially-
CC addressable set of single exon nucleic acid probes for measuring human
CC gene expression (comprising a plurality of single exon nucleic acid
CC probes cited above, where each of the plurality of probes is separately
CC and addressably isolatable or amplifiable from the plurality), a single
CC exon microarray for measuring human gene expression, a method of
CC measuring human gene expression, a vector comprising the single exon
CC probe cited above, an ORF-encoded peptide comprising at least 8
CC contiguous amino acids of any of the above-mentioned amino acid
CC sequences (optionally with conservative amino acid substitutions), an
CC isolated antibody that binds specifically to a peptide cited above,
XX
CC methods of selling and/or licensing single exon probes or microarrays to
CC a customer desiring to measure gene expression, a method of providing
CC human gene expression data by subscription, and a computer-readable
CC storage medium which contains a database having a plurality of records
CC (each record including data on the expression of a single exon probe
CC cited above. The probe, methods and apparatus are useful in gene
CC expression analysis. The probes may be used as tools for surveying
CC tissues to detect the presence of expressed messages that contain their
CC specific exon, or in constructing genome-derived single exon microarrays.
CC In addition, the probes are used in identifying and characterizing
CC alternative splicing events, in detecting and characterizing gross
CC alterations in the genomic locus that includes their exon, in assessing
CC smaller genomic alterations, in priming the synthesis of nucleic acids,
CC or in expressing the ORF-encoded peptide. The present sequence is a human
CC single exon probe of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20030194704
XX
SQ Sequence 491 BP; 95 A; 163 C; 145 G; 88 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 3.67 Length: 491
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 31.03% Indels: 0
DB: 12 Gaps: 0
US-10-624-218-2 (1-29) x ACH88921 (1-491)
Qy 1 ArgAlaAlaPheLeuGluLysGluAsn 9
Db 95 CGGGGGGCTTCCTGGAGAGAGAGAAC 121
RESULT 12
ACH75221
ID ACH75221 standard; DNA; 546 BP.
XX
AC ACH75221;
XX
DT 29-JUL-2004 (first entry)
XX
DE Human genome derived single exon probe #8416.
XX
KW Human; probe; ss; gene expression; single exon probe; microarray;
KW alternative splicing event; genomic alteration.
XX
OS Homo sapiens.
XX
PN US2003194704-A1.
XX
PD 16-OCT-2003.
XX
PF 03-APR-2002; 2002US-00029386.
XX
PR 03-APR-2002; 2002US-00029386.
XX
PA (PENN/) PENN S G.
PA (RANK/) RANK D R.
PA (HANZ/) HANZEL D K.
XX
PI Penn SG, Rank DR, Hanzel DK;
XX
DR WPI; 2004-119264/12.
XX
PT New human genome-derived single exon nucleic acid probes useful for human
PT gene expression analysis, for identifying or characterizing alternative
PT splicing events, for assessing genomic alterations or as tools for
PT surveying tissues.
XX
PS Claim 15; SEQ ID NO 8416; 80pp; English.
XX

CC The invention relates to a nucleic acid probe for measuring human gene
CC expression, comprising any of the 27,400 fully defined nucleotide
CC sequences in the specification, or their complements or fragments, and
CC encoding at least 8 amino acids of any of the 6888 amino acid sequences
CC fully defined in the specification. The probe is a single exon probe that
CC hybridises under high stringency conditions to a nucleic acid molecule
CC expressed in human cells or tissues. Also included are a spatially-
CC addressable set of single exon nucleic acid probes for measuring human
CC gene expression (comprising a plurality of single exon nucleic acid
CC probes cited above, where each of the plurality of probes is separately
CC and addressably isolatable or amplifiable from the plurality), a single
CC exon microarray for measuring human gene expression, a method of
CC measuring human gene expression, a vector comprising the single exon
CC probe cited above, an ORF-encoded peptide comprising at least 8
CC contiguous amino acids of any of the above-mentioned amino acid
CC sequences (optionally with conservative amino acid substitutions), an
CC isolated antibody that binds specifically to a peptide cited above,
CC a customer desiring to measure gene expression, a method of providing
CC human gene expression data by subscription, and a computer-readable
CC storage medium which contains a database having a plurality of records
CC (each record including data on the expression of a single exon probe
CC cited above). The probe, methods and apparatus are useful in gene
CC expression analysis. The probes may be used as tools for surveying
CC tissues to detect the presence of expressed messages that contain their
CC specific exon, or in constructing genome-derived single exon microarrays.
CC In addition, the probes are used in identifying and characterising gross
CC alternative splicing events, in detecting and characterising gross
CC alterations in the genomic locus that includes their exon, in assessing
CC smaller genomic alterations, in priming the synthesis of nucleic acids,
CC or in expressing the ORF-encoded peptide. The present sequence is a human
CC single exon probe of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20030194704

XX Sequence 546 BP; 106 A; 179 C; 161 G; 100 T; 0 U; 0 Other;

Alignment Scores: 4.05 Length: 546
Pred. No.: 9.00 Matches: 9
Score: 100.00% Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 31.03% Gaps: 0
DB: 12

US-10-624-218-2 (1-29) x ACH75221 (1-546)

QY 1 ArgAlaAlaPheLeuGluLysGluAsn 9
Db 120 CGGGCGGCTTCTCGAGAGGAGAAC 146

RESULT 13
ID ACH77332/c
XX ACH77332 standard; DNA; 546 BP.
AC ACH77332;
XX
XX 29-JUL-2004 (first entry)
DT
XX Human genome derived single exon probe #10527.
DE
XX Human; probe; ss; gene expression; single exon probe; microarray;
KW alternative splicing event; genomic alteration.
XX Homo sapiens.

OS
XX US2003194704-A1.
XX
XX 16-OCT-2003.
PD
XX 03-APR-2002; 2002US-00029386.
PF
XX

03-APR-2002; 2002US-00029386.

(PENN/) PENN S G.
(RANK/) RANK D R.
(HANZ/) HANZEL D K.

Penn SG, Rank DR, Hanzel DK;
WPI; 2004-119264/12.

New human genome-derived single exon nucleic acid probes useful for human
gene expression analysis, for identifying or characterizing alternative
splicing events, for assessing genomic alterations or as tools for
surveying tissues.

Claim 15; SEQ ID NO 10527; 80pp; English.

The invention relates to a nucleic acid probe for measuring human gene
expression, comprising any of the 27,400 fully defined nucleotide
sequences in the specification, or their complements or fragments, and
encoding at least 8 amino acids of any of the 6888 amino acid sequences
fully defined in the specification. The probe is a single exon probe that
hybridises under high stringency conditions to a nucleic acid molecule
expressed in human cells or tissues. Also included are a spatially-
addressable set of single exon nucleic acid probes for measuring human
gene expression (comprising a plurality of single exon nucleic acid
probes cited above, where each of the plurality of probes is separately
and addressably isolatable or amplifiable from the plurality), a single
exon microarray for measuring human gene expression, a method of
measuring human gene expression, a vector comprising the single exon
probe cited above, an ORF-encoded peptide comprising at least 8
contiguous amino acids of any of the above-mentioned amino acid
sequences (optionally with conservative amino acid substitutions), an
isolated antibody that binds specifically to a peptide cited above,
methods of selling and/or licensing single exon probes or microarrays to
a customer desiring to measure gene expression, a method of providing
human gene expression data by subscription, and a computer-readable
storage medium which contains a database having a plurality of records
(each record including data on the expression of a single exon probe
cited above). The probe, methods and apparatus are useful in gene
expression analysis. The probes may be used as tools for surveying
tissues to detect the presence of expressed messages that contain their
specific exon, or in constructing genome-derived single exon microarrays.
In addition, the probes are used in identifying and characterising gross
alternative splicing events, in detecting and characterising gross
alterations in the genomic locus that includes their exon, in assessing
smaller genomic alterations, in priming the synthesis of nucleic acids,
or in expressing the ORF-encoded peptide. The present sequence is a human
single exon probe of the invention. Note: The sequence data for this
patent did not form part of the printed specification, but was obtained
in electronic format directly from USPTO at
seqdata.uspto.gov/sequence.html?DocID=20030194704

Sequence 546 BP; 100 A; 161 C; 179 G; 106 T; 0 U; 0 Other;

Alignment Scores: 4.05 Length: 546
Pred. No.: 9.00 Matches: 9
Score: 100.00% Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 31.03% Gaps: 0
DB: 12

US-10-624-218-2 (1-29) x ACH77332 (1-546)

QY 1 ArgAlaAlaPheLeuGluLysGluAsn 9
Db 427 CGGGCGGCTTCTCGAGAGGAGAAC 401

RESULT 14
ADP13309
ID ADP13309 standard; DNA; 1385 BP.

XX

```
AC ADP13309;
XX
XX 26-AUG-2004 (first entry)
XX
XX Renal cell carcinoma differentially expressed gene #45.
DE
XX ds; diagnosis; non-blood disease; solid tumor; gene expression;
KW peripheral blood mononuclear cell; renal cell carcinoma; prostate cancer;
KW head/neck cancer; differential expression.
XX
XX Homo sapiens.
XX
XX WO2004048933-A2.
XX
XX 10-JUN-2004.
XX
XX 21-NOV-2003; 2003WO-US037481.
XX
XX 21-NOV-2002; 2002US-0427982P.
XX
XX 03-APR-2003; 2003US-0459782P.
XX
XX (AMHP ) WYETH.
XX
XX (TWIN/) TWINE N C.
XX
XX (BURC/) BURCZYNSKI M E.
XX
XX (TREP/) TREPICCHIO W L.
XX
XX (DORN/) DORNER A.
XX
XX (STOV/) STOVER J A.
XX
XX (SLON/) SLONI D K.
XX
XX Twine NC, Burczynski ME, Trepicchio WL, Dorner A, Stover JA;
PI Sloni DK;
PI
XX
XX WPI; 2004-460799/43.
XX
XX Diagnosing non-blood disease such as solid tumor, involves comparing
PT differential expression profile of specific genes in peripheral blood
PT sample of subject with reference expression profile of specific genes.
XX
XX Disclosure; SEQ ID NO 45; 350pp; English.
XX
XX The invention relate to a method of diagnosing (M1) non-blood disease
CC such as solid tumor by providing peripheral blood sample of human having
CC non-blood disease, and comparing an expression profile of specific genes
CC in the peripheral blood sample to reference expression profile of the
CC genes, where each of the genes is differentially expressed in peripheral
CC blood mononuclear cells (PBMCs) of patients having the disease as
CC compared to PBMCs of normal humans. The method is useful for diagnosing
CC non-blood disease such as solid tumor. The solid tumor is chosen from
CC renal cell carcinoma (RCC), prostate cancer and head/neck cancer. The
CC peripheral blood sample comprises enriched PBMCs. The peripheral blood
CC sample is a whole blood sample (claimed). (M1) is useful for identifying
CC genes that are differentially expressed in peripheral blood samples
CC isolated at different stages of progression, development or treatment of
CC RCC and/or other solid tumors. This sequence corresponds to a gene that
CC is differentially expressed and detected by the method of the invention.
CC (Note: this sequence is not given as part of the printed specification
CC but was obtained from WIPO in electronic format at
CC ftp.wipo.int/pub/published_pct_sequences/).
XX
XX Sequence 1385 BP; 279 A; 445 C; 404 G; 254 T; 0 U; 3 Other;
SQ
Alignment Scores:
Pred. No.: 9.56 Length: 1385
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 31.03% Indels: 0
DB: 12 Gaps: 0
US-10-624-218-2 (1-29) x ADP13309 (1-1385)
Qy 1 ArgAlaAlaPheLeuGluIysGluAsn 9
|||||
```

```
Db 255 CGGGCGGCTTCCTGGAGAGGAGAAC 281
RESULT 15
ADR52783
ID ADR52783 standard; DNA; 1385 BP.
XX
XX ADR52783;
AC
XX 18-NOV-2004 (first entry)
DT
XX Drug therapy altered expressed gene #134.
DE
XX drug activity monitoring; expression profile; gene expression;
KW peripheral blood sample; peripheral blood mononuclear cell; drug therapy;
KW CCI-779; immunosuppressant; rapamycin; mammalian target of rapamycin;
KW mTOR; ds.
XX
XX Homo sapiens.
XX
XX WO2004072265-A2.
XX
XX 26-AUG-2004.
XX
XX 11-FEB-2004; 2004WO-US004118.
XX
XX 11-FEB-2003; 2003US-0446133P.
XX
XX 03-APR-2003; 2003US-0459782P.
XX
XX 23-JAN-2004; 2004US-0538246P.
XX
XX (AMHP ) WYETH.
XX
XX (BURC/) BURCZYNSKI M.
XX
XX (TWIN/) TWINE N.
XX
XX (DORN/) DORNER A J.
XX
XX (TREP/) TREPICCHIO W L.
XX
XX Burczynski M, Twine N, Dorner AJ, Trepicchio WL;
PI WPI; 2004-642301/62.
XX
XX Monitoring drug activities in vivo comprises comparing an expression
PT profile of a gene in a peripheral blood sample of a patient before and
PT after drug therapy.
XX
XX Disclosure; SEQ ID NO 134; 136pp; English.
XX
XX The invention relates to a method of monitoring drug activities in vivo
CC by comparing an expression profile of at least one gene in a peripheral
CC blood sample of a patient to a reference expression profile of the at
CC least one gene, where the at least one gene is differentially expressed
CC in peripheral blood mononuclear cells (PBMCs) of patients who have a non-
CC blood disease and are subjected to a drug therapy as compared to PBMCs
CC isolated from the patient before the drug therapy, and where the patient
CC has the non-blood disease and is being treated by the drug therapy. The
CC method, kit, and nucleic acid array are useful for monitoring drug
CC activities in vivo. The drug is especially CCI-779, an ester analogue of
CC the immunosuppressant rapamycin which is a potent inhibitor of the
CC mammalian target of rapamycin (mTOR). This sequence represents a gene
CC expressed in PBMC altered by the drug therapy. (Note: this sequence does
CC no form part of the printed specification but was obtained in electronic
CC format from WIPO at ftp.wipo.int/pub/published_pct_sequences/).
XX
XX Sequence 1385 BP; 279 A; 445 C; 404 G; 254 T; 0 U; 3 Other;
SQ
Alignment Scores:
Pred. No.: 9.56 Length: 1385
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 31.03% Indels: 0
DB: 13 Gaps: 0
US-10-624-218-2 (1-29) x ADR52783 (1-1385)
```

Qy 1 ArgAlaAlaPheIeuGluLysGluAsn 9
|||
Db 255 CGGGCGGCTTCCTGGAGAGAGAAC 281

Search completed: July 27, 2005, 12:14:30
Job time : 236.403 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 27, 2005, 11:44:30 ; Search time 67.6667 Seconds
(without alignments)
701.261 Million cell updates/sec

Title: US-10-624-218-2

Perfect score: 29

Sequence: 1 RAAFLEKENTALRTVAELRKVGRCRN1 29

Scoring table:

OLIGO 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Word size: 1

Total number of hits satisfying chosen parameters: 2398959

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters: -DEV=xlp

-MODE=frame+ p2n model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US10624218/runat_26072005_121436_3711/app_query.fasta_1.796
-DB=Issued_Patents_NA -OPMT=fastap -SUFFIX=olip2n.rn1 -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -SPRT=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10624218 @CEN 1 1 213 @runat_26072005_121436_3711 -NCPU=6 -ICPU=3
-NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THRSADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA:
1: /cgn2_6/ptodata/1/ina/5A COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	14	48.3	1165	4	US-09-949-016-5024
2	14	48.3	32584	4	Sequence 5024, Ap
3	9	31.0	1370	4	Sequence 16766, A
4	9	31.0	10350	4	Sequence 5339, Ap
5	8	27.6	25	4	Sequence 17081, A
6	8	27.6	601	4	Sequence 19607-96363
7	8	27.6	194933	4	Sequence 40823, A
8	8	27.6	462589	4	Sequence 14172, A
9	8	27.6	476044	4	Sequence 12900, A
10	7	24.1	25	4	Sequence 12412, A
11	7	24.1	197	4	Sequence 96362, A
12	7	24.1	255	4	Sequence 21922, A
					Sequence 9464, Ap

13	7	24.1	264	4	US-09-134-000C-257	Sequence 257, App
C 14	7	24.1	457	4	US-09-270-767-2475	Sequence 2475, Ap
C 15	7	24.1	457	4	US-09-270-767-17757	Sequence 17757, A
C 16	7	24.1	600	4	US-09-902-540-5548	Sequence 5548, Ap
C 17	7	24.1	601	4	US-09-949-016-23956	Sequence 23956, A
C 18	7	24.1	601	4	US-09-949-016-148643	Sequence 148643, A
C 19	7	24.1	601	4	US-09-949-016-195386	Sequence 195386, A
C 20	7	24.1	606	4	US-09-902-540-1288	Sequence 1288, Ap
C 21	7	24.1	846	4	US-09-902-540-7051	Sequence 7051, Ap
C 22	7	24.1	1045	3	US-09-221-017B-953	Sequence 953, App
C 23	7	24.1	1128	4	US-09-498-520A-23	Sequence 23, Appl
C 24	7	24.1	1643	4	US-09-902-540-6808	Sequence 6808, Ap
C 25	7	24.1	2757	3	US-09-221-017B-1105	Sequence 1105, Ap
C 26	7	24.1	4282	4	US-09-902-540-563	Sequence 563, App
C 27	7	24.1	4835	4	US-09-902-540-627	Sequence 627, App
C 28	7	24.1	5006	4	US-09-620-312D-310	Sequence 310, App
C 29	7	24.1	14634	4	US-09-949-016-17384	Sequence 17384, A
C 30	7	24.1	15095	4	US-09-902-540-1077	Sequence 1077, Ap
C 31	7	24.1	16782	4	US-09-949-016-17291	Sequence 17291, A
C 32	7	24.1	20116	4	US-09-949-016-16861	Sequence 16861, A
C 33	7	24.1	22906	4	US-09-949-016-16471	Sequence 16471, A
C 34	7	24.1	26257	4	US-09-949-016-16791	Sequence 16791, A
C 35	7	24.1	28809	4	US-09-949-016-15825	Sequence 15825, A
C 36	7	24.1	39113	4	US-09-949-016-15634	Sequence 15634, A
C 37	7	24.1	40546	4	US-09-949-016-12847	Sequence 12847, A
C 38	7	24.1	40546	4	US-09-949-016-12915	Sequence 12915, A
C 39	7	24.1	40655	4	US-09-949-016-12032	Sequence 12032, A
C 40	7	24.1	40655	4	US-09-949-016-15919	Sequence 15919, A
C 41	7	24.1	41170	4	US-09-902-540-1367	Sequence 1267, Ap
C 42	7	24.1	42000	4	US-10-081-563-25	Sequence 25, Appl
C 43	7	24.1	47471	4	US-09-949-016-12271	Sequence 12271, A
C 44	7	24.1	107140	4	US-09-949-016-14834	Sequence 14834, A
C 45	7	24.1	113283	4	US-09-949-016-16976	Sequence 16976, A

ALIGNMENTS

RESULT 1

US-09-949-016-5024
; Sequence 5024, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949, 016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5024
; LENGTH: 1165
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-5024

Alignment Scores:
Pred. No.: 1.61e-05 Length: 1165
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 48.28% Indels: 0
DB: 4 Gaps: 0

US-10-624-218-2 (1-29) x US-09-949-016-5024 (1-1165)

QY 1 ArgAlaAlaPheLeuGluLysGluAsnThrAlaLeuArgThr 14

Db	991	CGGGCAGCCTTCTCTGGAGAGGAGACACACAGCCTTGGGAGC	1032
RESULT 2			
US-09-949-016-16766			
; Sequence 16766, Application US/09949016			
; Patent No. 6812339			
; GENERAL INFORMATION:			
; APPLICANT: VENTER, J. Craig et al.			
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED			
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF			
; FILE REFERENCE: CL001307			
; CURRENT APPLICATION NUMBER: US/09/949,016			
; CURRENT FILING DATE: 2000-04-14			
; PRIOR FILING DATE: 2000-10-20			
; PRIOR FILING DATE: 2000-10-20			
; PRIOR FILING DATE: 2000-10-03			
; PRIOR FILING DATE: 2000-10-03			
; PRIOR FILING DATE: 2000-09-08			
; NUMBER OF SEQ ID NOS: 207012			
; SOFTWARE: FastSEQ for Windows Version 4.0			
; SEQ ID NO 16766			
; LENGTH: 32584			
; TYPE: DNA			
; ORGANISM: Human			
US-09-949-016-16766			
Alignment Scores:			
Pred. No.:	0.000344	Length:	32584
Score:	14.00	Matches:	14
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	48.28%	Indels:	0
DB:	4	Gaps:	0
US-10-624-218-2 (1-29) x US-09-949-016-16766 (1-32584)			
QY	1	ArgAlaAlaPheLeuGluLysGluAsnThrAlaLeuArgThr	14
Db	30410	CGGGCAGCCTTCTCTGGAGAGGAGACACACAGCCTTGGGAGC	30451
RESULT 3			
US-09-949-016-5339			
; Sequence 5339, Application US/09949016			
; Patent No. 6812339			
; GENERAL INFORMATION:			
; APPLICANT: VENTER, J. Craig et al.			
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED			
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF			
; FILE REFERENCE: CL001307			
; CURRENT APPLICATION NUMBER: US/09/949,016			
; CURRENT FILING DATE: 2000-04-14			
; PRIOR APPLICATION NUMBER: 60/241,755			
; PRIOR FILING DATE: 2000-10-20			
; PRIOR APPLICATION NUMBER: 60/237,768			
; PRIOR FILING DATE: 2000-10-03			
; PRIOR APPLICATION NUMBER: 60/231,498			
; PRIOR FILING DATE: 2000-09-08			
; NUMBER OF SEQ ID NOS: 207012			
; SOFTWARE: FastSEQ for Windows Version 4.0			
; SEQ ID NO 5339			
; LENGTH: 1370			
; TYPE: DNA			
; ORGANISM: Human			
US-09-949-016-5339			
Alignment Scores:			
Pred. No.:	2.45	Length:	1370
Score:	9.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	31.03%	Indels:	0
US-10-624-218-2 (1-29) x US-09-949-016-5339			
QY	1	ArgAlaAlaPheLeuGluLysGluAsn	9
Db	8154	CGGGCGGCCTTCTCTGGAGAGGAGAAC	8180
RESULT 4			
US-09-949-016-17081			
; Sequence 17081, Application US/09949016			
; Patent No. 6812339			
; GENERAL INFORMATION:			
; APPLICANT: VENTER, J. Craig et al.			
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED			
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF			
; FILE REFERENCE: CL001307			
; CURRENT APPLICATION NUMBER: US/09/949,016			
; CURRENT FILING DATE: 2000-04-14			
; PRIOR APPLICATION NUMBER: 60/241,755			
; PRIOR FILING DATE: 2000-10-20			
; PRIOR APPLICATION NUMBER: 60/237,768			
; PRIOR FILING DATE: 2000-10-03			
; PRIOR APPLICATION NUMBER: 60/231,498			
; NUMBER OF SEQ ID NOS: 207012			
; SOFTWARE: FastSEQ for Windows Version 4.0			
; SEQ ID NO 17081			
; LENGTH: 10550			
; TYPE: DNA			
; ORGANISM: Human			
US-09-949-016-17081			
Alignment Scores:			
Pred. No.:	16	Length:	10550
Score:	9.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	31.03%	Indels:	0
DB:	4	Gaps:	0
US-10-624-218-2 (1-29) x US-09-949-016-17081 (1-10550)			
QY	1	ArgAlaAlaPheLeuGluLysGluAsn	9
Db	8154	CGGGCGGCCTTCTCTGGAGAGGAGAAC	8180
RESULT 5			
US-09-396-196G-96363			
; Sequence 96363, Application US/09396196G			
; Patent No. 6821724			
; GENERAL INFORMATION:			
; APPLICANT: Michael Mittmann			
; APPLICANT: David Mack			
; APPLICANT: David Lockhart			
; APPLICANT: Affymetrix, Inc.			
; TITLE OF INVENTION: Methods of Genetic Analysis			
; FILE REFERENCE: 3101.1			
; CURRENT APPLICATION NUMBER: US/09/396,196G			
; CURRENT FILING DATE: 1999-09-15			
; PRIOR APPLICATION NUMBER: 60/100,678			
; PRIOR FILING DATE: 1998-09-17			
; NUMBER OF SEQ ID NOS: 127806			
; SOFTWARE: FastSEQ for Windows Version 4.0			
; SEQ ID NO 96363			
; LENGTH: 25			
; TYPE: DNA			
; ORGANISM: mus musculus			
US-09-396-196G-96363			
Alignment Scores:			
Pred. No.:	0.655	Length:	25
Score:	8.00	Matches:	8

Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 27.59% Indels: 0
 DB: 4 Gaps: 0

US-10-624-218-2 (1-29) x US-09-396-196G-96363 (1-25)

Oy 3 AlaPheLeuGluLysGluAsnThr 10
 Db 2 GCGTTCCTGGAGAGGAGACACA 25

RESULT 6

US-09-949-016-40823
 ; Sequence 40823, Application US/09949016
 ; Patent No. 6812339

GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

FILE REFERENCE: CL001307

CURRENT APPLICATION NUMBER: US/09/949,016

PRIOR FILING DATE: 2000-04-14

PRIOR APPLICATION NUMBER: 60/241,755

PRIOR FILING DATE: 2000-10-20

PRIOR APPLICATION NUMBER: 60/237,768

PRIOR FILING DATE: 2000-10-03

PRIOR APPLICATION NUMBER: 60/231,498

PRIOR FILING DATE: 2000-09-08

NUMBER OF SEQ ID NOS: 207012

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 40823

LENGTH: 601

TYPE: DNA

ORGANISM: Human

US-09-949-016-40823

Alignment Scores:

Pred. No.: 12.1 Length: 601
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 27.59% Indels: 0
 DB: 4 Gaps: 0

US-10-624-218-2 (1-29) x US-09-949-016-40823 (1-601)

Oy 4 PheLeuGluLysGluAsnThrAla 11
 Db 32 TTTTGTAGAAAAGAAACACTGCT 55

RESULT 7

US-09-949-016-14172/c
 ; Sequence 14172, Application US/09949016
 ; Patent No. 6812339

GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

FILE REFERENCE: CL001307

CURRENT APPLICATION NUMBER: US/09/949,016

PRIOR FILING DATE: 2000-04-14

PRIOR APPLICATION NUMBER: 60/241,755

PRIOR FILING DATE: 2000-10-20

PRIOR APPLICATION NUMBER: 60/237,768

PRIOR FILING DATE: 2000-10-03

PRIOR APPLICATION NUMBER: 60/231,498

PRIOR FILING DATE: 2000-09-08

NUMBER OF SEQ ID NOS: 207012

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 14172

LENGTH: 194933

TYPE: DNA

ORGANISM: Human

FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)...(194933)
 ; OTHER INFORMATION: n = A,T,C or G
 US-09-949-016-14172

Alignment Scores:

Pred. No.: 2.46e+03 Length: 194933
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 27.59% Indels: 0
 DB: 4 Gaps: 0

US-10-624-218-2 (1-29) x US-09-949-016-14172 (1-194933)

Oy 15 ArgValAlaGluLeuArgLysArg 22

Db 32921 AGAGTGGCTGAATTGAGGAAAAGA 32898

RESULT 8

US-09-949-016-12900/c

; Sequence 12900, Application US/09949016

; Patent No. 6812339

GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

FILE REFERENCE: CL001307

CURRENT APPLICATION NUMBER: US/09/949,016

PRIOR FILING DATE: 2000-04-14

PRIOR APPLICATION NUMBER: 60/241,755

PRIOR FILING DATE: 2000-10-20

PRIOR APPLICATION NUMBER: 60/237,768

PRIOR FILING DATE: 2000-10-03

PRIOR APPLICATION NUMBER: 60/231,498

PRIOR FILING DATE: 2000-09-08

NUMBER OF SEQ ID NOS: 207012

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 12900

LENGTH: 462589

TYPE: DNA

ORGANISM: Human

US-09-949-016-12900

Alignment Scores:

Pred. No.: 5.43e+03 Length: 462589
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 27.59% Indels: 0
 DB: 4 Gaps: 0

US-10-624-218-2 (1-29) x US-09-949-016-12900 (1-462589)

Oy 4 PheLeuGluLysGluAsnThrAla 11

Db 281446 TTTTGTAGAAAAGAAACACTGCT 281423

RESULT 9

US-09-949-016-12412/c

; Sequence 12412, Application US/09949016

; Patent No. 6812339

GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

FILE REFERENCE: CL001307

CURRENT APPLICATION NUMBER: US/09/949,016

PRIOR FILING DATE: 2000-04-14

PRIOR APPLICATION NUMBER: 60/241,755

PRIOR FILING DATE: 2000-10-20

PRIOR APPLICATION NUMBER: 60/237,768

```

; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12412
; LENGTH: 476044
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12412

Alignment Scores:
Pred. No.: 5.57e+03 Length: 476044
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 27.59% Indels: 0
DB: 4 Gaps: 0

US-10-624-218-2 (1-29) x US-09-949-016-12412 (1-476044)

Qy 4 PheLeuGluLysGluAsnThrAla 11
Db 294902 TTTTGGAGAAAGAAACACTGCT 294879

RESULT 10
US-09-396-196G-96362
; Sequence 96362, Application US/09396196G
; Patent No. 6821724
; GENERAL INFORMATION:
; APPLICANT: Michael Mittmann
; APPLICANT: David Mack
; APPLICANT: David Lockhart
; APPLICANT: Affymetrix, Inc.
; TITLE OF INVENTION: Methods of Genetic Analysis
; FILE REFERENCE: 3101.1
; CURRENT APPLICATION NUMBER: US/09/396,196G
; CURRENT FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: 60/100,678
; PRIOR FILING DATE: 1998-09-17
; NUMBER OF SEQ ID NOS: 127806
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 96362
; LENGTH: 25
; TYPE: DNA
; ORGANISM: mus musculus
US-09-396-196G-96362

Alignment Scores:
Pred. No.: 6.91 Length: 25
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 24.14% Indels: 0
DB: 4 Gaps: 0

US-10-624-218-2 (1-29) x US-09-396-196G-96362 (1-25)

Qy 1 ArgAlaAlaPheLeuGluLys 7
Db 5 CGGGCAGCGTTCCTGGAGAAG 25

RESULT 11
US-09-513-999C-21922
; Sequence 21922, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG

; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 21922
; LENGTH: 197
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-513-999C-21922

Alignment Scores:
Pred. No.: 46 Length: 197
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 24.14% Indels: 0
DB: 4 Gaps: 0

US-10-624-218-2 (1-29) x US-09-513-999C-21922 (1-197)

Qy 4 PheLeuGluLysGluAsnThr 10
Db 86 TTTTGGAGAAAGAGAATACA 106

RESULT 12
US-09-902-540-9464
; Sequence 9464, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 9464
; LENGTH: 255
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-9464

Alignment Scores:
Pred. No.: 58.3 Length: 255
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 24.14% Indels: 0
DB: 4 Gaps: 0

US-10-624-218-2 (1-29) x US-09-902-540-9464 (1-255)

Qy 1 ArgAlaAlaPheLeuGluLys 7
Db 208 CGGGCGGCGCTTCCTTGAAGAAG 228

RESULT 13
US-09-134-000C-257
; Sequence 257, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
```


; PRIOR APPLICATION NUMBER: US 60/055,778
 ; PRIOR FILING DATE: 1997-08-15
 ; NUMBER OF SEQ ID NOS: 6812
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 257
 ; LENGTH: 264
 ; TYPE: DNA
 ; ORGANISM: Enterococcus faecalis
 US-09-134-000C-257

Alignment Scores:
 Pred. No.: 60.2 Length: 264
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 24.14% Indels: 0
 DB: 4 Gaps: 0

US-10-624-218-2 (1-29) x US-09-134-000C-257 (1-264)

Qy 3 AlaPheLeuGluLysGluAsn 9
 |||||
 Db 53 GCCTTTITAGAGAAAGAAAT 73

RESULT 14

US-09-270-767-2475/c
 ; Sequence 2475, Application US/09270767
 ; Patent No. 6703491
 ; GENERAL INFORMATION:
 ; APPLICANT: Homburger et al.
 ; TITLE OF INVENTION: Nucleic acids and proteins of *Drosophila melanogaster*
 ; FILE REFERENCE: File Reference: 7326-094
 ; CURRENT APPLICATION NUMBER: US/09/270,767
 ; CURRENT FILING DATE: 1999-03-17
 ; NUMBER OF SEQ ID NOS: 62517
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 2475
 ; LENGTH: 457
 ; TYPE: DNA
 ; ORGANISM: *Drosophila melanogaster*
 US-09-270-767-2475

Alignment Scores:
 Pred. No.: 99.7 Length: 457
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 24.14% Indels: 0
 DB: 4 Gaps: 0

US-10-624-218-2 (1-29) x US-09-270-767-2475 (1-457)

Qy 16 ValAlaGluLeuArgLysArg 22
 |||||
 Db 28 GTGGCTGAGCTGCGAAAGCGG 8

RESULT 15

US-09-270-767-17757/c
 ; Sequence 17757, Application US/09270767
 ; Patent No. 6703491
 ; GENERAL INFORMATION:
 ; APPLICANT: Homburger et al.
 ; TITLE OF INVENTION: Nucleic acids and proteins of *Drosophila melanogaster*
 ; FILE REFERENCE: File Reference: 7326-094
 ; CURRENT APPLICATION NUMBER: US/09/270,767
 ; CURRENT FILING DATE: 1999-03-17
 ; NUMBER OF SEQ ID NOS: 62517
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 17757
 ; LENGTH: 457
 ; TYPE: DNA
 ; ORGANISM: *Drosophila melanogaster*
 US-09-270-767-17757

Alignment Scores:
 Pred. No.: 99.7 Length: 457
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 24.14% Indels: 0
 DB: 4 Gaps: 0

US-10-624-218-2 (1-29) x US-09-270-767-17757 (1-457)

Qy 16 ValAlaGluLeuArgLysArg 22
 |||||
 Db 28 GTGGCTGAGCTGCGAAAGCGG 8

Search completed: July 27, 2005, 16:49:10
 Job time : 142.667 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 27, 2005, 11:55:09 ; Search time 368.34 Seconds
(without alignments)
509.100 Million cell updates/sec

Title: US-10-624-218-2

Perfect score: 29
Sequence: 1 RAAPLEXENTALTRVAELKRVGRCRNI 29

Scoring table:
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 7277826 seqs, 3233139505 residues

Word size: 1
Total number of hits satisfying chosen parameters: 14543013

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:
-MODE=frame_p2n.model -DEV=xlp
-Q=/cgn2_1/USPRO_spool/p/US10624218/runat_26072005_121437_3743/app_query.fasta_1.796
-DB=Published Applications NA -QMT=fastap -SUFFIX=olip2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=quality -THR MIN=1
-ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10624218@cgn 1.1.1189 @runat_26072005_121437_3743
-NCPUE=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=60 -YGAPEXT=6 -DELOP=6 -DELEXT=7

Database : Published Applications NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
19: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq.*
20: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq.*
21: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq.*
22: /cgn2_6/ptodata/1/pubpna/US10J_PUBCOMB.seq.*
23: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
24: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
25: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
26: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	9	31.0	217	16	US-10-029-386-24227
C 2	9	31.0	466	10	US-09-918-995-23172
C 3	9	31.0	491	16	US-10-029-386-22116
C 4	9	31.0	546	16	US-10-029-386-8416
C 5	9	31.0	546	16	US-10-029-386-10527
C 6	9	31.0	1385	19	US-10-717-597-45
C 7	9	31.0	1385	19	US-10-775-169-134
C 8	9	31.0	1478	21	US-10-887-553A-202
C 9	9	31.0	1638	21	US-10-764-420-2185
C 10	9	31.0	1671	17	US-10-388-934-173
C 11	9	31.0	1671	21	US-10-870-387-20
C 12	8	27.6	25	21	US-10-719-900-429122
C 13	8	27.6	25	21	US-10-719-900-429125
C 14	8	27.6	25	21	US-10-809-189-96363
C 15	8	27.6	1365	9	US-09-738-626-2418
C 16	8	27.6	1517	19	US-10-781-014-563
C 17	8	27.6	2025	9	US-09-738-626-2417
C 18	8	27.6	2155	21	US-10-494-836-115
C 19	8	27.6	3000	15	US-10-156-761-5949
C 20	8	27.6	58215	13	US-10-087-192-379
C 21	8	27.6	96594	11	US-09-997-722-22
C 22	8	27.6	101505	13	US-10-087-192-1270
C 23	8	27.6	3309400	9	US-09-738-626-1
C 24	8	27.6	9025608	15	US-10-156-761-1
C 25	7	24.1	25	21	US-10-809-189-96362
C 26	7	24.1	201	19	US-10-741-601-2446
C 27	7	24.1	202	19	US-10-674-124A-3492
C 28	7	24.1	290	14	US-10-060-036-1337
C 29	7	24.1	315	19	US-10-430-201-3839
C 30	7	24.1	326	17	US-10-012-697-676
C 31	7	24.1	343	19	US-10-674-124A-9837
C 32	7	24.1	381	11	US-09-864-408A-2041
C 33	7	24.1	382	19	US-10-674-124A-1899
C 34	7	24.1	409	20	US-10-425-115-78684
C 35	7	24.1	424	20	US-10-425-115-52879
C 36	7	24.1	424	20	US-10-425-115-43863
C 37	7	24.1	450	19	US-10-437-963-24197
C 38	7	24.1	475	19	US-10-767-701-26967
C 39	7	24.1	477	10	US-09-918-995-16654
C 40	7	24.1	536	22	US-10-972-079-93228
C 41	7	24.1	573	21	US-10-487-901-2105
C 42	7	24.1	585	13	US-10-027-632-234817
C 43	7	24.1	585	17	US-10-027-632-234817
C 44	7	24.1	591	20	US-10-363-345A-17549
C 45	7	24.1			

ALIGNMENTS

RESULT 1
US-10-029-386-24227/c
; Sequence 24227, Application US/10029386
; Publication NO. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 24227

```

; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 22116
; LENGTH: 491
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR19.3
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.7
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.79
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.3
; OTHER INFORMATION: SWISSPROT HIT: Q10586, EVALUE 4.00e-34
; OTHER INFORMATION: NT HIT: D28468.1, EVALUE 0.00e+00
; OTHER INFORMATION: EST_HUMAN HIT: A1684518.1, EVALUE 0.00e+00
US-10-029-386-22116

Alignment Scores:
Pred. No.: 3.01 Length: 491
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 31.03% Indels: 0
DB: 16 Gaps: 0

US-10-624-218-2 (1-29) x US-10-029-386-22116 (1-491)
QY 1 ArgAlaAlaPheLeuGluLysGluAsn 9
DB 95 CGGGCGGCCTTCCTGGAGAGGAGAAC 121

RESULT 4
US-10-029-386-8416
; Sequence 8416, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 8416
; LENGTH: 546
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR19.3
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.7
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.79
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.3
; OTHER INFORMATION: SWISSPROT HIT: Q10586, EVALUE 4.00e-34
; OTHER INFORMATION: NT HIT: U48213.1, EVALUE 0.00e+00
; OTHER INFORMATION: EST_HUMAN HIT: A1684518.1, EVALUE 0.00e+00
US-10-029-386-8416

Alignment Scores:
Pred. No.: 3.3 Length: 546
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
DB: 16 Gaps: 0

US-10-624-218-2 (1-29) x US-10-029-386-24227 (1-217)
QY 1 ArgAlaAlaPheLeuGluLysGluAsn 9
DB 123 CGGGCGGCCTTCCTGGAGAGGAGAAC 97

RESULT 2
US-09-918-995-23172
; Sequence 23172, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 23172
; LENGTH: 466
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-918-995-23172

Alignment Scores:
Pred. No.: 2.87 Length: 466
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 31.03% Indels: 0
DB: 10 Gaps: 0

US-10-624-218-2 (1-29) x US-09-918-995-23172 (1-466)
QY 1 ArgAlaAlaPheLeuGluLysGluAsn 9
DB 407 CGGGCGGCCTTCCTGGAGAGGAGAAC 433

RESULT 3
US-10-029-386-22116
; Sequence 22116, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 22116
; LENGTH: 491
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR19.3
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.7
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.79
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.3
; OTHER INFORMATION: SWISSPROT HIT: Q10586, EVALUE 4.00e-34
; OTHER INFORMATION: NT HIT: D28468.1, EVALUE 0.00e+00
; OTHER INFORMATION: EST_HUMAN HIT: A1684518.1, EVALUE 0.00e+00
US-10-029-386-24227

Alignment Scores:
Pred. No.: 1.48 Length: 217
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 31.03% Indels: 0
DB: 16 Gaps: 0
```

Query Match: 31.03% Indels: 0
 DB: 16 Gaps: 0

US-10-624-218-2 (1-29) x US-10-029-386-8416 (1-546)

Qy 1 ArgAlaAlaPheLeuGluLysGluAsn 9
 Db 120 CGGGCGGCTTCTCGAGAGGAGAAC 146

RESULT 5

US-10-029-386-10527/c
 ; Sequence 10527, Application US/10029386
 ; Publication No. US20030194704A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Penn, Sharron G.
 ; APPLICANT: Rank, David R.
 ; APPLICANT: Hanzel, David K.
 ; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR C
 ; FILE REFERENCE: AEWICA-X-2
 ; CURRENT APPLICATION NUMBER: US/10/029,386
 ; CURRENT FILING DATE: 2001-12-20
 ; NUMBER OF SEQ ID NOS: 34288
 ; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
 ; SEQ ID NO 10527
 ; LENGTH: 546
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION: MAP TO CHR19.3
 ; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.6
 ; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.2
 ; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.4
 ; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 8.7
 ; OTHER INFORMATION: NT HIT: U49213.1, EVALUE 0.00e+00
 ; OTHER INFORMATION: EST HUMAN HIT: A1684518.1, EVALUE 0.00e+00
 ; OTHER INFORMATION: SWISSPROT HIT: Q10586, EVALUE 4.00e-34

US-10-029-386-10527

Alignment Scores: 3.3 Length: 546
 Pred. No.: 9.00 Matches: 9
 Score: 100.00% Conservative: 0
 Percent Similarity: 100.00% Mismatches: 0
 Best Local Similarity: 100.00% Indels: 0
 Query Match: 31.03% Gaps: 0
 DB: 16

US-10-624-218-2 (1-29) x US-10-029-386-10527 (1-546)

Qy 1 ArgAlaAlaPheLeuGluLysGluAsn 9
 Db 427 CGGGCGGCTTCTCGAGAGGAGAAC 401

RESULT 6

US-10-717-597-45
 ; Sequence 45, Application US/10717597
 ; Publication No. US20040110221A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wyeth
 ; APPLICANT: Burczynski, Michael E.
 ; APPLICANT: Twine, Natalie C.
 ; APPLICANT: Dorner, Andrew J.
 ; APPLICANT: Trepicchio, William L.
 ; APPLICANT: Slonim, Donna K.
 ; APPLICANT: Stover, Jennifer A.
 ; TITLE OF INVENTION: METHODS FOR DIAGNOSING RCC AND OTHER SOLID TUMORS
 ; FILE REFERENCE: AM101080L
 ; CURRENT APPLICATION NUMBER: US/10/717,597
 ; CURRENT FILING DATE: 2003-11-21
 ; PRIOR APPLICATION NUMBER: US 60/459,782
 ; PRIOR FILING DATE: 2003-04-03
 ; PRIOR APPLICATION NUMBER: US 60/427,982
 ; PRIOR FILING DATE: 2002-11-21

NUMBER OF SEQ ID NOS: 4904
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 45
 ; LENGTH: 1385
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (20)..(1158)
 ; OTHER INFORMATION: The "n" residue at position 1158 is nucleotide g according to the
 ; OTHER INFORMATION: corresponding sequence (DBP gene) in the Entrez Human Genome
 ; OTHER INFORMATION: Sequence Database. No residue exists at n position 20 and 25
 ; OTHER INFORMATION: according to the same corresponding sequence.

US-10-717-597-45

Alignment Scores: 7.42 Length: 1385
 Pred. No.: 9.00 Matches: 9
 Score: 100.00% Conservative: 0
 Percent Similarity: 100.00% Mismatches: 0
 Best Local Similarity: 100.00% Indels: 0
 Query Match: 31.03% Gaps: 0
 DB: 19

US-10-624-218-2 (1-29) x US-10-717-597-45 (1-1385)

Qy 1 ArgAlaAlaPheLeuGluLysGluAsn 9
 Db 255 CGGGCGGCTTCTCGAGAGGAGAAC 281

RESULT 7

US-10-775-169-134
 ; Sequence 134, Application US/10775169
 ; Publication No. US20040175743A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wyeth
 ; APPLICANT: Burczynski, Michael
 ; APPLICANT: Twine, Natalie
 ; APPLICANT: Dorner, Andrew
 ; APPLICANT: Trepicchio, William
 ; TITLE OF INVENTION: Method for Monitoring Drug Activities In Vivo
 ; FILE REFERENCE: AM101080 (031896-013000)
 ; CURRENT APPLICATION NUMBER: US/10/775,169
 ; CURRENT FILING DATE: 2004-02-11
 ; NUMBER OF SEQ ID NOS: 5278
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 134
 ; LENGTH: 1385
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (20)..(1158)
 ; OTHER INFORMATION: Each "n" represents a nucleotide selected from a, t, g or c, or
 ; OTHER INFORMATION: contains no nucleotide.

US-10-775-169-134

Alignment Scores: 7.42 Length: 1385
 Pred. No.: 9.00 Matches: 9
 Score: 100.00% Conservative: 0
 Percent Similarity: 100.00% Mismatches: 0
 Best Local Similarity: 100.00% Indels: 0
 Query Match: 31.03% Gaps: 0
 DB: 19

US-10-624-218-2 (1-29) x US-10-775-169-134 (1-1385)

Qy 1 ArgAlaAlaPheLeuGluLysGluAsn 9
 Db 255 CGGGCGGCTTCTCGAGAGGAGAAC 281

RESULT 8

US-10-887-553A-202
 ; Sequence 202, Application US/10887553A


```
RESULT 12
US-10-719-900-429122
; Sequence 429122, Application US/10719900
; Publication No. US20050026164A1
; GENERAL INFORMATION:
; APPLICANT: Xue Mei Zhou
; TITLE OF INVENTION: Methods of Genetic Analysis of Mouse
; FILE REFERENCE: 3528.1
; CURRENT APPLICATION NUMBER: US/10/719,900
; PRIOR FILING DATE: 2003-11-20
; PRIOR APPLICATION NUMBER: 60/427,808
; PRIOR FILING DATE: 2002-11-20
; NUMBER OF SEQ ID NOS: 982914
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 429122
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-719-900-429122

Alignment Scores:
Pred. No.: 2.53 Length: 25
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 27.59% Indels: 0
DB: 21 Gaps: 0

US-10-624-218-2 (1-29) x US-10-719-900-429122 (1-25)

QY 7 LysGluAanThrAlaLeuArgThr 14
| | | | | | | | | | | | | | | | | | | | |
DB 2 AAGGAGAACACTGCGCGGACG 25

RESULT 13
US-10-719-900-429125
; Sequence 429125, Application US/10719900
; Publication No. US20050026164A1
; GENERAL INFORMATION:
; APPLICANT: Xue Mei Zhou
; TITLE OF INVENTION: Methods of Genetic Analysis of Mouse
; FILE REFERENCE: 3528.1
; CURRENT APPLICATION NUMBER: US/10/719,900
; PRIOR FILING DATE: 2003-11-20
; PRIOR APPLICATION NUMBER: 60/427,808
; PRIOR FILING DATE: 2002-11-20
; NUMBER OF SEQ ID NOS: 982914
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 429125
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-719-900-429125

Alignment Scores:
Pred. No.: 2.53 Length: 25
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 27.59% Indels: 0
DB: 21 Gaps: 0

US-10-624-218-2 (1-29) x US-10-719-900-429125 (1-25)

QY 7 LysGluAanThrAlaLeuArgThr 14
| | | | | | | | | | | | | | | | | | | | |
DB 2 AAGGAGAACACTGCGCGGACG 25

RESULT 14
US-10-809-189-96363
; Sequence 96363, Application US/10809189
; Publication No. US20050048531A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Michael Mittmann
; APPLICANT: David Mack
; APPLICANT: David Lockhart
; APPLICANT: Affymetrix, Inc.
; TITLE OF INVENTION: Methods of Genetic Analysis
; FILE REFERENCE: 3101.1
; CURRENT APPLICATION NUMBER: US/10/809,189
; CURRENT FILING DATE: 2004-03-25
; PRIOR APPLICATION NUMBER: US/09/396,196
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: 60/100,678
; PRIOR FILING DATE: 1998-09-17
; NUMBER OF SEQ ID NOS: 127806
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 96363
; LENGTH: 25
; TYPE: DNA
; ORGANISM: mus musculus
US-10-809-189-96363

Alignment Scores:
Pred. No.: 2.53 Length: 25
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 27.59% Indels: 0
DB: 21 Gaps: 0

US-10-624-218-2 (1-29) x US-10-809-189-96363 (1-25)

QY 3 AlapheteuGluLysGluAanThr 10
| | | | | | | | | | | | | | | | | | | | |
DB 2 CGGTTCTCGGAGAGGAGACACA 25

RESULT 15
US-09-738-626-2418/c
; Sequence 2418, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OKAZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 2418
; LENGTH: 1365
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-2418

Alignment Scores:
Pred. No.: 82.6 Length: 1365
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 27.59% Indels: 0
```

DB: 9 Gaps: 0

US-10-624-218-2 (1-29) x US-09-738-626-2418 (1-1365)

Qy 14 ThrArgValAlaGluLeuArgLys 21

Db 117 ACTGCGTTGCTGAGCTGCGTAAG 94

Search completed: July 27, 2005, 17:17:46
Job time : 376.34 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 27, 2005, 11:14:04 ; Search time 2016.31 Seconds
(without alignments)
547.469 Million cell updates/sec

Title: US-10-624-218-2

Perfect score: 29
Sequence: 1 RAAPFLEKENTALRTVAELKRVGRCRNI 29

Scoring table:
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Word size: 1
Total number of hits satisfying chosen parameters: 68476236

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlp
-Q=/cgm2_1/USPRO_pool_PUS10624218/runat_26072005_121436_3698/app_query.fasta_1.796
-DB=EST -QFMT=fastap -SUFFIX=olip2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10624218 @CGN 1.1 9525 @runat_26072005_121436_3698 -NCPU=6 -ICPU=3
-NO MMAP -LARGQUERY -NEG_SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : EST:
1: gb_est1.*
2: gb_est2.*
3: gb_hic.*
4: gb_est3.*
5: gb_est4.*
6: gb_est5.*
7: gb_est6.*
8: gb_gss1.*
9: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	14	48.3	341	1	AI014348
2	14	48.3	351	6	BY775604
3	14	48.3	380	1	AI892971
4	14	48.3	398	6	CB707074
5	14	48.3	418	7	CN394275
6	14	48.3	440	2	BE295612
7	14	48.3	446	1	AA138848
8	14	48.3	464	4	BG986696
9	14	48.3	484	5	BU262464

10	14	48.3	528	2	BF5660216
11	14	48.3	540	7	CR536622
12	14	48.3	555	2	AW836791
13	14	48.3	557	1	BU111019
14	14	48.3	578	1	AL120537
15	14	48.3	599	2	AW500333
16	14	48.3	626	5	BU469699
17	14	48.3	631	6	CB577428
18	14	48.3	638	7	CK624751
19	14	48.3	654	4	BG432348
20	14	48.3	658	9	CR828685
21	14	48.3	673	5	BM963436
22	14	48.3	693	6	CB247229
23	14	48.3	786	5	BU466941
24	14	48.3	788	7	CN455515
25	14	48.3	812	5	BQ042190
26	14	48.3	850	2	BF309490
27	14	48.3	928	6	CB590113
28	14	48.3	933	5	BQ887766
29	14	48.3	941	2	BE737088
30	14	48.3	958	5	BU149688
31	13	44.8	520	2	AW637693
32	13	44.8	591	4	BJ095767
33	13	44.8	694	5	BX852163
34	13	44.8	707	5	BP702661
35	13	44.8	709	5	BX846930
36	13	44.8	919	5	BU899556
37	13	44.8	922	5	BQ733819
38	11	37.9	563	7	CO794673
39	11	37.9	653	2	BF203280
40	11	37.9	846	5	BU441005
41	10	34.5	498	2	AW836776
42	10	34.5	1034	4	BM466864
43	9	31.0	268	7	F26608
44	9	31.0	287	4	BM709261
45	9	31.0	289	4	BM693068

ALIGNMENTS

RESULT 1

AI014348

LOCUS

DEFINITION

am5gil.s1 Johnston frontal cortex Homo sapiens cDNA clone

(IMAGE:1539140 3', similar to gb:M95586 TRANSCRIPTION FACTOR E2-ALPHA

(HUMAN); mRNA sequence.

ACCESSION

AI014348

VERSION

AI014348.1

KEYWORDS

EST.

SOURCE

Homo sapiens

ORGANISM

Homo sapiens

REFERENCE

AUTHORS

Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S.,

Krizman, D., Kucaba, T., Lacy, M., Lennon, G., Marra, M.,

Martin, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F.,

Theising, B., White, Y., Wylie, T., Waterston, R. and Wilson, R.

WashU-NCI human EST Project

Unpublished (1997)

CONTACT: Wilton RK

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@wustl.wustl.edu

This clone is available royalty-free through LNL ; contact the

BF5660216	maa29a04.
CR536622	DKF2p459A
AW836791	QV1-LT003
BU111019	603540871
AL120537	DKF2p761A
AW500333	UI-HF-BN0
BU469699	603761076
CB577428	AMGNNUC.N
CK624751	mi24909.Y
BG432348	602496669
CR828685	GROAAA60A
BM963436	UI-M-EQO-
CB247229	UI-M-FIO-
BU466941	603371693
CN455515	UI-M-HNO-
BQ042190	UI-M-EQO-
BF309490	601892191
CB590113	AGENCOURT
BQ887766	AGENCOURT
BE737088	601304854
BU149688	AGENCOURT
AW637693	bl1e01.w
BJ095767	BJ095767
BX852163	BX852163
BP702661	BP702661
BX846930	BX846930
BU899556	AGENCOURT
BQ733819	AGENCOURT
CO794673	NT144B.B0
BF203280	601866155
BU441005	604146587
AW836776	QV1-LT003
BM466864	AGENCOURT
F26608	HSPD14153.H
BM709261	UI-B-CQ1-
BM693068	UI-B-CQ1-

```

/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1539140"
/sex="male"
/tissue_type="pooled frontal lobe"
/dev_stage="adult"
/lab_host="SOLR (kanamycin resistant)"
/clone_lib="Johnston frontal cortex"
/note="Organ: brain; Vector: Bluescript SK-; Site_1:
ECORI; Stanley Neuropathology Consortium
(www.stanleylab.org) brains S-58, S-65, S-67, S-78.
Random + oligo-dT primed into EcoRI site of ZAP II Vector.
Mass excised. Avg insert length 1.9kb. Custom library
provided by Dr. Nancy Johnston [(410) 614-3918,
nlj@welchlink.welch.jhu.edu].

ORIGIN
Alignment Scores:      Length:      341
Pred. No.:      0.000112      Matches:      14
Score:      14.00      Conservaive:      0
Percent Similarity:      100.00%      Mismatches:      0
Best Local Similarity:      100.00%      Indels:      0
Query Match:      48.28%      Gaps:      0
DB:      1

US-10-624-218-2 (1-29) x AI014348 (1-341)

Qy      1 ArgAlaAlaPheLeuGluLysGluAsnThrAlaLeuArgThr 14
Db      131 CGGGCAGCCTTCCTGGAGAGGAGACACACGCCCTGGCAGC 172

RESULT 2
BY775604
LOCUS
DEFINITION
Mus musculus cDNA clone L930096D17 5', mRNA sequence.
VERSION
BY775604.1 GI:39702242
KEYWORDS
EST.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
REFERENCE
AUTHORS
1 (bases 1 to 351)
Alzawa,K., Arakawa,T., Ishii,Y., Sasaki,D., Bono,H., Kondo,S.,
Sugahara,Y., Saito,R., Osato,N., Fukuda,S., Sato,K., Watanaki,A.,
Hirozane-Kishikawa,T., Nakamura,M., Shibata,Y., Yasunishi,A.,
Kikuchi,N., Yoshiki,A., Kusakabe,M., Gustincich,S., Beisel,K.,
Pavan,W., Aidinis,V., Nakagawara,A., Held,W.A., Iwata,H., Kono,T.,
Nakauchi,H., Lyons,P., Wells,C., Hume,D.A., Fagioliini,M.,
Hensch,T.K., Brinkmeier,M., Camper,S., Hirota,J., Mombaerts,P.,
Muramatsu,M., Okazaki,Y., Kawai,J. and Hayashizaki,Y.
Targeting a complex transcriptome: the construction of the mouse
full-length cDNA encyclopedia
Genome Res. 13 (6B), 1273-1289 (2003)
22703353
12819125
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Sushiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site (http://genome.gsc.riken.jp/) for
further details.

FEATURES
source
Location/Qualifiers
1..351
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="L930096D17"
/tissue_type="whole body"
/dev_stage="17.5 days embryo"
/clone_lib="RIKEN full-length enriched, 17.5 days embryo
whole body"

ORIGIN
Alignment Scores:      Length:      351
Pred. No.:      0.000115      Matches:      14
Score:      14.00      Conservaive:      0
Percent Similarity:      100.00%      Mismatches:      0
Best Local Similarity:      100.00%      Indels:      0
Query Match:      48.28%      Gaps:      0
DB:      0

US-10-624-218-2 (1-29) x BY775604 (1-351)

Qy      1 ArgAlaAlaPheLeuGluLysGluAsnThrAlaLeuArgThr 14
Db      24 CGGGCAGCCTTCCTGGAGAGGAGACACACGCCCTGGCAGC 65

RESULT 3
AI892971
LOCUS
DEFINITION
mq82h05.y1 Stratagene mouse melanoma (#937312) Mus musculus cDNA
clone IMAGE:585273 5', similar to gb:M95586 TRANSCRIPTION FACTOR
E2-ALPHA (HUMAN); , mRNA sequence.
ACCESSION
AI892971
VERSION
AI892971.1 GI:5598873
KEYWORDS
EST.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
REFERENCE
AUTHORS
1 (bases 1 to 380)
Marra,M., Hillier,B., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y.,
Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R.,
Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
Waterston,R. and Wilson,R.
The WashU-NCI Mouse EST Project 1999
Unpublished (1999)
Contact: Marra M/WashU-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
This read is a RESEQUENCE of a previously sequenced mouse clone
This read has been verified (found to hit its original self in the
correct orientation)
MGI:359921
Seq primer: -40RP from Gibco
High quality sequence stop: 336.
Location/Qualifiers
1..380
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:585273"
/tissue_type="melanoma"
/dev_stage="M2 cells"
/lab_host="SOLR (kanamycin resistant)"
/clone_lib="Stratagene mouse melanoma (#937312)"
/note="Organ: skin; Vector: pBluescript SK-; Site_1:

```

EcoRI, Site 2: XhoI; Cloned unidirectionally. Primer:
Oligo dT. From M2 cells, a highly metastatic derivative of
the K-1735 (mouse) melanoma. Average insert size: 1.0 kb;
Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGGCAGCAG
3' -3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'

ORIGIN

Alignment Scores:
Pred. No.: 0.000124 Length: 380
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservativeness: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 48.28% Indels: 0
DB: 1 Gaps: 0

US-10-624-218-2 (1-29) x AI892971 (1-380)

Qy 1 ArgAlaAlaPheLeuGluLysGluAsnThrAlaLeuArgThr 14
Db 89 CGGGCAGCCTTCTGGAGAAGGAGAACACAGCCCTGCGGACG 130

RESULT 4

CB707074 398 bp mRNA linear EST 10-APR-2003
LOCUS AMGNNUC:TRCP2-00001-D6-A trcp2 (10289) Rattus norvegicus cDNA clone
DEFINITION trcp2-00001-d6 5', mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Rattus norvegicus (Norway rat)
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE

1 (bases 1 to 398)
Amgen EST Program.

AUTHORS

TITLE

JOURNAL

COMMENT

Contact: Dan Fitzpatrick
Amgen, Inc
One Amgen Center Drive, Thousand Oaks, CA 91320-1799, USA
Tel: 805 447-4881

Plate: 00001 row: d column: 6.

FEATURES

source

1..398
Location/Qualifiers

/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/clone="trcp2-00001-d6"
/tissue type="choroid plexus brain"
/clone lib="trcp2 (10289)"
/note="Vector: C6KGF7L; Site_1: SalI; Site_2: NotI;
choroid plexus brain region"

ORIGIN

Alignment Scores:
Pred. No.: 0.000129 Length: 398
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservativeness: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 48.28% Indels: 0
DB: 6 Gaps: 0

US-10-624-218-2 (1-29) x CB707074 (1-398)

Qy 1 ArgAlaAlaPheLeuGluLysGluAsnThrAlaLeuArgThr 14
Db 50 CGGGCAGCCTTCTGGAGAAGGAGAACACAGCCCTGCGGACG 91

RESULT 5

CB707074 418 bp mRNA linear EST 16-MAY-2004
LOCUS CN394275
DEFINITION 1700045536536 GRN_ES Homo sapiens cDNA 5', mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1 (bases 1 to 418)
Brandenberger, R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G.J.,
Li, Y., Xu, C., Fang, R., Guegler, K., Rao, M.S., Mandalam, R.,
Lebkowski, J. and Stanton, L.W.
Transcriptome characterization elucidates signaling networks that
control human ES cell growth and differentiation
Nat. Biotechnol. 22 (6), 707-716 (2004)
Contact: Brandenberger R
Regenerative Medicine
Genon Corporation
230 Constitution Drive, Menlo Park, CA 94025, USA
Tel: 650 473 8658
Fax: 650 473 7760
Email: rbrandenberger@genon.com
Insert Length: 418 Std Error: 0.00.

FEATURES

source

1..418
Location/Qualifiers

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="embryonic stem cells, cell lines H1, H7, and
H9"
/clone lib="GRN ES"
/note="oligo dT-primed, full-length enriched cDNA library
from undifferentiated hES cell lines H1 (p32), H7 (p29),
and H9 (p26) maintained in feeder-free conditions"

ORIGIN

Alignment Scores:

Pred. No.: 0.000135 Length: 418

Score: 14.00 Matches: 14

Percent Similarity: 100.00% Conservativeness: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 48.28% Indels: 0

DB: 7 Gaps: 0

US-10-624-218-2 (1-29) x CN394275 (1-418)

Qy

1 ArgAlaAlaPheLeuGluLysGluAsnThrAlaLeuArgThr 14

Db 68 CGGGCAGCCTTCTGGAGAAGGAGAACACAGCCCTGCGGACG 109

RESULT 6

BE295612

LOCUS BE295612

DEFINITION 601176577F1 NTH_MGC_17 Homo sapiens cDNA clone IMAGE:3531762 5',

mRNA sequence.

Accession BE295612.1 GI:9179163

Version BE295612

Keywords EST.

Source Homo sapiens (human)

Organism Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 440)

NIH-MGC http://mgi.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgabbs@mail.nih.gov

Tissue Procurement: ATCC

cDNA Library Prepared by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov

Plate: L1CM206 row: j column: 19
 High quality sequence start: 6
 High quality sequence stop: 437.
 Location/Qualifiers

FEATURES

source

1. .440
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:3531762"
 /tissue_type="rhabdomyosarcoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NH.MGC.17"
 /note="Organ: muscle; Vector: pOTB7; Site 1: EcoRI; Site 2: XhoI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

ORIGIN

Alignment Scores:
 Pred. No.: 0.000142 Length: 440
 Score: 14.00 Matches: 14
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 48.28% Indels: 0
 Gaps: 2

US-10-624-218-2 (1-29) x BE295612 (1-440)

QY 1 ArgAlaAlaPheLeuGluLysGluAsnThrAlaLeuArgThr 14
 DB 327 CGGCAGCGCTTCCTGGAGAGGAGAACACACGCCCTGCGGAGC 368

RESULT 7

AA138848 446 bp mRNA linear EST 09-FEB-1997
 LOCUS mq92h05.r1 Stratagene mouse melanoma (#937312) Mus musculus cDNA
 DEFINITION clone IMAGE:585273 5' similar to gb:M95586 TRANSCRIPTION FACTOR
 E2-ALPHA (HUMAN) mRNA sequence.

ACCESSION AA138848
 VERSION AA138848.1 GI:1701049
 KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 446)
 Marra M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisler, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M., Scheinberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wyllie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R.

TITLE The WashU-HMI Mouse EST Project
 JOURNAL Unpublished (1996)
 COMMENT Contact: Marra M/Mouse EST Project

WashU-HMI Mouse EST Project
 Washington University School of MedicineP
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810

Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LBNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:359921
 Seq primer: -28m13 rev1 ET from Amersham
 High quality sequence stop: 443.

FEATURES

source

1. .446
 /organism="Mus musculus"
 /mol_type="mRNA"

/db_xref="taxon:10090"
 /clone="IMAGE:585273"
 /tissue_type="melanoma"
 /dev_stage="M2 cells"
 /lab_host="SOLR (kanamycin resistant)"
 /clone_lib="Stratagene mouse melanoma (#937312)"
 /note="Organ: skin; Vector: pBluescript SK-; Site 1: EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer: Oligo dT. From M2 cells, a highly metastatic derivative of the K-1735 (mouse) melanoma. Average insert size: 1.0 kb; Uni-ZAP XR Vector; ~5' adaptor sequence: 5' GAATTTCGACGACGAG 3' ~3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3"

ORIGIN

Alignment Scores:
 Pred. No.: 0.000144 Length: 446
 Score: 14.00 Matches: 14
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 48.28% Indels: 0
 Gaps: 1

US-10-624-218-2 (1-29) x AA138848 (1-446)

QY 1 ArgAlaAlaPheLeuGluLysGluAsnThrAlaLeuArgThr 14
 DB 88 CGGCAGCGCTTCCTGGAGAGGAGAACACACGCCCTGCGGAGC 129

RESULT 8

EG986696 464 bp mRNA linear EST 13-JUN-2001
 LOCUS PM1-HT1170-030101-001-b09 HT1170 Homo sapiens cDNA, mRNA sequence.

ACCESSION EG986696
 VERSION EG986696.1 GI:14390766

KEYWORDS

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 464)

REFERENCE

AUTHORS

Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H., Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.

Shotgun sequencing of the human transcriptome with ORF expressed

sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

JOURNAL

MEDLINE

PUBMED

COMMENT

Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=PM1&t2=PM1-HT1170-030101-001-b09&t3=2001-01-03&t4=1)

Seq primer: puc 18 forward

High quality sequence start: 6

High quality sequence stop: 396.

FEATURES

source

1. .464
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /dev_stage="Adult"
 /clone_lib="HT1170"

/note="Organ: head neck; Vector: puc19; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN

Alignment Scores:
Pred. No.: 0.000149 Length: 464
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 48.28% Indels: 0
DB: 4 Gaps: 0

US-10-624-218-2 (1-29) x BG986696 (1-464)

QY 1 ArgAlaAlaPheLeuGluLysGluAsnThrAlaLeuArgThr 14
|||||
DB 169 CGGCAGCCTTCCTGGAGAAGAGAACACAGCCTCGCGACG 210

RESULT 9

BU262464 603502757F1 CSEQCHN51 Gallus gallus mRNA linear EST 26-NOV-2002
LOCUS 484 bp mRNA clone CHEST422b16 5', mRNA
DEFINITION sequence.

ACCESSION

BU262464

VERSION BU262464.1 GI:25530860

KEYWORDS EST.

SOURCE Gallus gallus (chicken)

ORGANISM Gallus gallus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Archosauria; Aves; Neognathae; Galliformes; Phasianidae;

Phasianinae; Gallus.

REFERENCE 1 (bases 1 to 484)

AUTHORS Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E.,

Fong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.

A Comprehensive Collection of Chicken cDNAs

Curr. Biol. 12 (22), 1965-1969 (2002)

MEDLINE 22335534

PUBMED 12445392

COMMENT Contact: Simon Hubbard

Department of Biomolecular Sciences

University of Manchester Institute of Science and Technology

(UMIST)

PO Box 88, Manchester, M60 1QD, UK

Tel: 01612008930

Fax: 01612360409

Email: Simon.Hubbard@umist.ac.uk.

Location/Qualifiers

1. .484

/organism="Gallus gallus"

/mol_type="mRNA"

/strain="White Leghorn, Hisex"

/db_xref="taxon:9031"

/clone="CHEST422b16"

/dev_stage="22"

/lab_host="DH10B"

/clone_lib="CSEQCHN51"

/note="Organ: limbs; Vector: pBluescript II KS(+); Site_1:

EcoRI; Site_2: NotI; This normalized library was

constructed from 1 million independent clones. cDNA

synthesis was initiated using an oligo(dT) primer, using

methylated C in the first strand synthesis reaction.

Following this first strand reaction, double-stranded cDNA

was blunt-ended, ligated to NotI adapters, digested with

EcoRI, size-selected, and cloned into the NotI and EcoRI

compatible sites of a custom modified MCS of the

pBluescript (KS+) vector. The library was normalized in 2

rounds using conditions adapted from Soares et al., PNAS

(1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6

(1996): 791, except that a significantly longer reannealing hybridization was used."

ORIGIN

Alignment Scores:
Pred. No.: 0.000155 Length: 484
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 48.28% Indels: 0
DB: 5 Gaps: 0

US-10-624-218-2 (1-29) x BU262464 (1-484)

QY 1 ArgAlaAlaPheLeuGluLysGluAsnThrAlaLeuArgThr 14
|||||
DB 261 CGGCAGCCTTCCTTGAGAAAGAGAAATACGCGCTCGAGACG 302

RESULT 10

BF660216

LOCUS

DEFINITION

maaz29a04.y1 NCI CGAP L110 Mus musculus cDNA clone IMAGE:3812262 5'

similar to SW:TEF_RAT_P41224 THYROTROPH EMBRYONIC FACTOR. [1] ;

mRNA sequence.

ACCESSION BF660216

VERSION BF660216.1 GI:11925350

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

REFERENCE 1 (bases 1 to 528)

AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

Other ESTs: maaz29a04.x1

Contact: Robert Strausberg, Ph.D.

Email: cgapsb-remail.nih.gov

Tissue Procurement: Jeffrey E. Green, M.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

MGI:1454374

Seq primer: -40RP from Gibco

High quality sequence stop: 435.

Location/Qualifiers

1. .528

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

/clone="IMAGE:3812262"

/sex="female"

/dev_stage="10 weeks"

/lab_host="DH10B (T1 phage-resistant)"

/clone_lib="NCI CGAP L110"

/note="Organ: liver; Vector: pCMV-SPORT6; Site_1: NotI;

Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.

Average insert size 1.6 kb. Library constructed by Life

Technologies."

ORIGIN

Alignment Scores:

Pred. No.: 0.000168 Length: 528

Score: 14.00 Matches: 14

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 48.28% Indels: 0

DB: 2 Gaps: 0

Fong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.
A Comprehensive Collection of Chicken cDNAs
Curr. Biol. 12 (22), 1965-1969 (2002)
2233534
MEDLINE
12445392
PUBMED
COMMENT

Contact: Simon Hubbard
Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology
(UMIST)
PO Box 88, Manchester, M60 1QD, UK
Tel: 01612008930
Fax: 01612360409
Email: Simon.Hubbard@umist.ac.uk.

FEATURES
source
1. 557
Location/Qualifiers
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="Compton Line 151"
/db_xref="taxon:9031"
/clone="ChEST509n17"
/sex="Female"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="CSEQCIN61"
/note="Organ: heart; Vector: pBluescript II KS(+); Site 1:
EcoRI; Site 2: NotI; This normalized library was
constructed from 1 million independent clones. cDNA
synthesis was initiated using an oligo(dT) primer, using
methylated C in the first strand synthesis reaction.
Following this first strand reaction, double-stranded cDNA
was blunted, ligated to NotI adapters, digested with
EcoRI, size-selected, and cloned into the NotI and EcoRI
compatible sites of a custom modified MCS of the
pBluescript (KS+) vector. The library was normalized in 2
rounds using conditions adapted from Soares et al., PNAS
(1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
(1996): 791, except that a significantly longer
reannealing hybridization was used."

ORIGIN
Alignment Scores:
Pred. No.: 0.000176 Length: 557
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 48.28% Indels: 0
DB: 5 Gaps: 0

US-10-624-218-2 (1-29) x BU311019 (1-557)
QY 1 ArgAlaAlaPheLeuGluLysGluAnThrAlaLeuArgThr 14
|||||
DB 117 CGGGCAGCCTTCTCTGAGAAAGAGATACGGCCTCGAGACG 158
|||||

RESULT 14
AL120537
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Blum, H., Bauersachs, S., Mewes, H.W., Gassenhuber, J. and Wiemann, S.
EST (Blum, et al.)
Unpublished (1999)
Contact: MIPS
MIPS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany

This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by LMU (Ludwig Maximilians University,
Munich/Germany) within the cDNA sequencing consortium of the German
Genome Project.
No sl sequence available.
This clone (DKFZp761A019) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES
source
1. 578
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp761A019"
/tissue_type="amygdala"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="761 (synonym: hamy2)"
/note="Vector: pSport1; Site 1: NotI; Site 2: SalI"

ORIGIN
Alignment Scores:
Pred. No.: 0.000183 Length: 578
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 48.28% Indels: 0
DB: 1 Gaps: 0

US-10-624-218-2 (1-29) x AL120537 (1-578)
QY 1 ArgAlaAlaPheLeuGluLysGluAnThrAlaLeuArgThr 14
|||||
DB 370 CGGGCAGCCTTCTCTGAGAGAGACACACGCCCTCGGACG 411
|||||

RESULT 15
AW500333
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Eco RI site shown at the beginning of the sequence.
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
cDNA Library Preparation: M.B. Soares Lab
cDNA Library Arrayed by: M.B. Soares Lab
DNA Sequencing by: M.B. Soares Lab
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Seq primer: M13 Forward.

FEATURES
source
1. 599
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3077127"
/tissue_type="lymph"
/cell_line="MGC85"

/lab_host="DH10B (LTI)"
/clone_lib="NIH_MGC_50"
/notes="Vector: pT773-Pac; Site 1: NotI; Site 2: Eco RI;
Constructed from size fractionated cytoplasmic mRNA
(3.5-4.4kb). Directionally cloned. Cells provided by
Louis M. Staudt, Ph.D. Library preparation by Maria de
Fatima Bonaldo, Ph.D. and M. Bento Soares, Ph.D."

ORIGIN

Alignment Scores:
Pred. No.: 0.000189 Length: 599
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 48.28% Indels: 0
DB: 2 Gaps: 0

US-10-624-218-2 (1-29) x AW500333 (1-599)

Qy 1 ArgAlaAlaPheLeuGluLysGluAsnThrAlaLeuArgThr 14
Db 166 CGGGCAGCCTTCCTGGAGAGGAGACACAGCCCTGCGGACG 207

Search completed: July 27, 2005, 16:41:10
Job time : 2018.31 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 26, 2005, 11:54:19 ; Search time 108.694 Seconds
(without alignments)
153.004 Million cell updates/sec

Title: US-10-624-218-3

Perfect score: 43

Sequence: 1 LEIAAFLERNTALETRVA.....QRVRLNRNSQYTRYGPL 43

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 2105692 seqs, 386760381 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	43	100.0	43	ADI34278	Zipper pe
2	8	18.6	520	ABB61499	Abb61499 Drosophil
3	8	18.6	539	ADC07982	Adc07982 Rice prot
4	8	18.6	1162	Aau36235	Aau36235 Pseudomon
5	8	18.6	1162	ABU38370	Abu38370 Protein e
6	8	18.6	1382	ADI45379	Adi45379 Rice isop
7	7	16.3	29	ADH48506	Adh48506 HLA-B*44/
8	7	16.3	9	ADI34277	Adi34277 Zipper pe
9	7	16.3	114	AAG90757	Aag90757 C glutami
10	7	16.3	186	ADM18946	Admi18946 Transposo
11	7	16.3	192	ABO64927	Aboc64927 Klebsiell
12	7	16.3	217	ABO82848	Aboc82848 Pseudomon
13	7	16.3	341	ADB74286	Adb74286 Mycobacte
14	7	16.3	465	ABB59427	Abbs59427 Drosophil
15	7	16.3	477	ABO71770	Aboc71770 Pseudomon
16	7	16.3	481	ADJ69603	Adj69603 Human hea
17	7	16.3	485	ABO76122	Aboc76122 Pseudomon
18	7	16.3	498	AAB79524	Aab79524 Corynebac
19	7	16.3	515	AAW56750	Aaw56750 Ecdysterol
20	7	16.3	515	AAE02830	Aae02830 HZSNPV ec
21	7	16.3	532	AAR27362	Aar27362 Sequence
22	7	16.3	558	AAR27363	Aar27363 Sequence
23	7	16.3	593	ABBS8343	Abbs8343 Drosophil
24	7	16.3	675	AAG92163	Aag92163 C glutami
25	7	16.3	675	ADD13431	Add13431 C. glutam

26	7	16.3	790	8	ADS43080	Adsa43080 Bacterial
27	7	16.3	793	7	ADE61200	Adsa61200 Human Pro
28	7	16.3	793	7	ADE61196	Adsa61196 Human Pro
29	7	16.3	793	8	ADQ30525	Adq30525 Pancreas
30	7	16.3	1093	8	ADS21852	Adsa21852 Bacterial
31	6	14.0	16	2	AAW69135	Aaw69135 Neuronal
32	6	14.0	25	4	AAW18600	Aaw18600 Peptide #
33	6	14.0	25	4	AAW70752	Aaw70752 Human bon
34	6	14.0	25	4	AAW58287	Aaw58287 Human bra
35	6	14.0	30	2	AAR28977	Aar28977 Selectin
36	6	14.0	31	2	AAW35662	Aaw35662 T cell pe
37	6	14.0	32	4	AAW61548	Aaw61548 Peptide W
38	6	14.0	33	5	ABB04571	Abb04571 Hepatitis
39	6	14.0	33	8	ADL97609	Adl97609 Protein e
40	6	14.0	36	3	AAW22035	Aaw22035 Peptide #
41	6	14.0	37	4	AAW61570	Aaw61570 Peptide W
42	6	14.0	43	8	ADI34279	Adi34279 Zipper pe
43	6	14.0	50	7	ADM26648	Adm26648 Hyperther
44	6	14.0	52	5	ABB08249	Abb08249 HA epitop
45	6	14.0	58	5	ABP09095	Abp09095 Human ORF

ALIGNMENTS

RESULT 1
ADI34278
ID ADI34278 standard; protein; 43 AA.
XX
AC ADI34278;
XX
DT 15-APR-2004 (first entry)
XX
DE Zipper peptide #3 for cross linking adenoviral ligands.
XX
KW adenovirus vector; Cytostatic; HER2/neu; CD40; tumor necrosis factor;
KW TNF.
XX
OS Unidentified.
XX
PN WO2004009133-A1.
XX
PD 29-JAN-2004.
XX
PF 22-JUL-2003; 2003WO-US022852.
XX
PR 22-JUL-2002; 2002US-037951P.
XX
PA (VECT-) VECTORLOGICS INC.
XX
PI Korokhov N, Mikhveva G;
XX
DR WPI; 2004-132871/13.
XX
PT Novel recombinant adenovirus having fiber protein modified by insertion
PT of first zipper peptide that can crosslink to second zipper peptide-
PT targeting ligand fusion protein, and binding between zipper peptides
PT targets vector to cell.
XX
PS Claim 2; SEQ ID NO 3; 54pp; English.

XX
The present invention relates to a targeted recombinant adenovirus vector. The invention is useful for expressing a heterologous protein chosen from a tumor associated antigen, HER2/neu and carcinoembryonic antigen, in a target cell e.g., CD40 + cell such as dendritic cells, is useful in a method of gene transfer to CD40 + cells, where the targeted adenovirus vector mediates transfer of the gene encoding heterologous protein to the cell such as the dendritic cell. The vector is useful in gene therapy techniques for treatment of tumors. Multivalent interaction or trimeric CD40L with CD40 receptors causes CD40 ligation, which then results in enhanced survival of these cells and secretion of cytokines such as interleukin (IL)-1, IL-6, IL-8, IL-10, IL-12, tumor necrosis factor (TNF)- α agr γ , macrophage inflammatory protein (MIP)-1 α and enzymes

AAU36235 standard; protein; 1162 AA.
AAU36235;
14-FEB-2002 (first entry)
Pseudomonas aeruginosa cellular proliferation protein #225.
Antisense; prokaryotic cellular proliferation protein; antibiotic;
antibacterial; drug design.
Pseudomonas aeruginosa.
WO200170955-A2.
27-SEP-2001.
21-MAR-2001; 2001WO-US009180.
21-MAR-2000; 2000US-0191078P.
23-MAY-2000; 2000US-0206848P.
26-MAY-2000; 2000US-0207727P.
23-OCT-2000; 2000US-0242578P.
27-NOV-2000; 2000US-0253625P.
22-DEC-2000; 2000US-0257931P.
16-FEB-2001; 2001US-0269308P.
(ELIT-) ELITRA PHARM INC.
Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
Yamamoto RT, Xu HH;
WPI: 2001-611495/70.
N-PSDB; AAS54094.
New polynucleotides for the identification and development of
antibiotics, comprise sequences of antisense nucleic acids.
Example 3; SEQ ID NO 11828; 511pp; English.
The invention relates to antisense inhibitors of genes essential to
prokaryotic cellular proliferation, their use in identifying the genes,
their use in the discovery of novel antibiotics, the essential genes
themselves and the encoded proteins. The prokaryotes used are Escherichia
coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
useful for the identification of potential new targets for antibiotic
development. The antisense nucleic acids can also be used to identify
proteins used in proliferation, to express these proteins, and to obtain
antibodies capable of binding to the expressed proteins. The proteins can
be used to screen compounds in rational drug discovery programmes. The
antisense nucleic acid sequence is also useful to screen for homologous
nucleic acids which are required for cell proliferation in a wide variety
of organisms. The present sequence represents an essential prokaryotic
cellular proliferation protein. Note: The sequence data for this patent
did not form part of the printed specification, but was obtained in
electronic format directly from WIPO at
ftp.wipo.int/pub/published_pct_sequences
Sequence 1162 AA;
Query Match 18.6%; Score 8; DB 4; Length 1162;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 14 ALETRVAE 21
| | | | |
Db 665 ALETRVAE 672
RESULT 5
ABU38370
ID ABU38370 standard; protein; 1162 AA.
ABU38370;
19-JUN-2003 (first entry)
Protein encoded by Prokaryotic essential gene #23897.
Antisense; prokaryotic essential gene; cell proliferation; drug design.
Pseudomonas aeruginosa.
WO200277183-A2.
03-OCT-2002.
21-MAR-2002; 2002WO-US009107.
21-MAR-2001; 2001US-00815242.
06-SEP-2001; 2001US-00948993.
25-OCT-2001; 2001US-0342923P.
08-FEB-2002; 2002US-00072851.
06-MAR-2002; 2002US-0362699P.
(ELIT-) ELITRA PHARM INC.
Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
WPI: 2003-029926/02.
N-PSDB; ACA42240.
New antisense nucleic acids, useful for identifying proteins or screening
for homologous nucleic acids required for cellular proliferation to
isolate candidate molecules for rational drug discovery programs.
Claim 25; SEQ ID NO 66294; 1766pp; English.
The invention relates to an isolated nucleic acid comprising any one of
the 6213 antisense sequences given in the specification where expression
of the nucleic acid inhibits proliferation of a cell. Also included are:
(1) a vector comprising a promoter operably linked to the nucleic acid
encoding a polypeptide whose expression is inhibited by the antisense
nucleic acid; (2) a host cell containing the vector; (3) an isolated
polypeptide or its fragment whose expression is inhibited by the
antisense nucleic acid; (4) an antibody capable of specifically binding
the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
proliferation or the activity of a gene in an operon required for
proliferation; (7) identifying a compound that influences the activity of
the gene product or that has an activity against a biological pathway
required for proliferation, or that inhibits cellular proliferation; (8)
identifying a gene required for cellular proliferation or the biological
pathway in which a proliferation-required gene or its gene product lies
or a gene on which the test compound that inhibits proliferation of an
organism acts; (9) manufacturing an antibiotic; (10) profiling a
compound's activity; (11) a culture comprising strains in which the gene
product is overexpressed or underexpressed; (12) determining the extent
to which each of the strains is present in a culture or collection of
strains; or (13) identifying the target of a compound that inhibits the
proliferation of an organism. The antisense nucleic acids are useful for
identifying proteins or screening for homologous nucleic acids required
for cellular proliferation to isolate candidate molecules for rational
drug discovery programs, or for screening homologous nucleic acids
required for proliferation in cells other than S. aureus, S. typhimurium,
K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
the target prokaryotic essential genes. Note: The sequence data for this
patent did not form part of the printed specification, but was obtained
in electronic format directly from WIPO at
ftp.wipo.int/pub/published_pct_sequences
Sequence 1162 AA;
Query Match 18.6%; Score 8; DB 6; Length 1162;
Best Local Similarity 100.0%; Pred. No. 16;

```
Matches      8;  Conservative      0;  Mismatches      0;  Indels      0;  Gaps      0;

Qy      14 ALETRVAE 21
        |||||
Db      665 ALETRVAE 672

RESULT 6
ADI45379 standard; protein; 1382 AA.
XX
AC      ADI45379;
XX
DT      22-APR-2004 (first entry)
XX
DE      Rice isoprenoid biosynthesis-associated protein #155.
XX
KW      Rice; isoprenoid biosynthesis; plant; isopentenyl diphosphate; IPP;
KW      dimethylallyl alcohol; DMAPP; short-chain plastid prenyltransferase;
KW      gibberellin; carotenoid; abscisic acid; tocopherol; plastoquinone;
KW      phyloquinone; mevalonate pathway; phytosterol; brassinosteroid;
KW      ubiquinone; monoterpene; sesquiterpene; protein prenylation; chlorophyll;
KW      haeme; yield.
XX
OS      Oryza sativa.
XX
PN      US2004010815-A1.
XX
PD      15-JAN-2004.
XX
PF      26-SEP-2002; 2002US-00259194.
XX
PR      26-SEP-2001; 2001US-0325277P.
PR      04-APR-2002; 2002US-0370620P.
PR      04-APR-2002; 2002US-0370743P.
XX
XX      (LANG/) LANGE B M.
XX      (GHAS/) GHASSEMIAN M.
XX      (BRIG/) BRIGGS S P.
XX      (COOP/) COOPER B.
XX      (GLAZ/) GLAZEBROOK J.
XX      (GOFF/) GOFF S A.
XX      (KATA/) KATAGIRI F.
XX      (KREP/) KREPS J.
XX      (MOUG/) MOUGHAMER T.
XX      (PROV/) PROVART N.
XX      (RICK/) RICHE D.
XX      (ZHUT/) ZHU T.
XX
XX      Lange BM, Ghassemian M, Briggs SP, Cooper B, Glazebrook J;
XX      Goff SA, Katagiri F, Kreps J, Moughamer T, Provart N, Ricke D;
XX      Zhu T;
XX
XX      WPI; 2004-090562/09.
XX      N-PSDB; ADI45378.
XX
XX      New isolated polynucleotides and polypeptides associated with isoprenoid
XX      synthesis in plants, useful for producing transgenic plants, for targeted
XX      gene disruption, as well as markers or probes.
XX
XX      Claim 4; SEQ ID NO 310; 117pp; English.
XX
XX      The invention relates to a polynucleotide (or its complement, protein
XX      encoding fragment or reverse complement), comprising a nucleotide
XX      sequence encoding a polypeptide comprising an amino acid sequence
XX      involved in or associated with the biosynthesis of isoprenoids in a rice
XX      plant. Also included are an isolated polypeptide involved in or
XX      associated with the biosynthesis of isoprenoids in a plant, an expression
XX      cassette comprising the polynucleotide, a host cell comprising the
XX      expression cassette, and a transgenic plant comprising the expression
XX      cassette. The polypeptides and polynucleotides include those associated
XX      with the biosynthesis of isopentenyl diphosphate (IPP) and dimethylallyl
XX      alcohol (DMAPP), the biosynthesis of short-chain plastid
```

```
CC      prenyltransferases, the biosynthesis of gibberellins, the biosynthesis of
CC      carotenoids and/or abscisic acids, the biosynthesis of tocopherols,
CC      plastoquinone and/or phyloquinone biosynthesis, the mevalonate pathway,
CC      phytosterol and brassinosteroid metabolism, biosynthesis of ubiquinone,
CC      biosynthesis of monoterpenes and sesquiterpenes, protein prenylation, and
CC      biosynthesis of chlorophyll or haeme. Also disclosed are banana, wheat
CC      and corn homologues of some of the rice polynucleotides. The
CC      polynucleotides are useful for producing transgenic plants, where the
CC      genome is augmented by a nucleic acid molecule of the invention, or in
CC      which the corresponding gene has been disrupted, e.g. to result in a
CC      loss, a decrease or an alteration in the function of the product encoded
CC      by the gene. The plants may also have increased yields and/or produce a
CC      better quality product than the corresponding wild-type plant. The
CC      nucleic acid molecules are useful for targeted gene disruption, as well
CC      as markers and probes. Note: The sequence data for this patent did not
CC      form part of the printed specification, but was obtained in electronic
CC      format directly from USPTO at
CC      seqdata.uspto.gov/sequence.html?DocID=20040010815. The present sequence
CC      represents a Rice isoprenoid biosynthesis-associated protein of the
CC      invention.
XX
SQ      Sequence 1382 AA;

Query Match      18.6%; Score 8; DB 8; Length 1382;
Best Local Similarity 100.0%; Fred. No. 19;
Matches      8;  Conservative      0;  Mismatches      0;  Indels      0;  Gaps      0;

Qy      25 RVQRLRNR 32
        |||||
Db      716 RVQRLRNR 723

RESULT 7
ADH48506 standard; peptide; 9 AA.
XX
AC      ADH48506;
XX
DT      25-MAR-2004 (first entry)
XX
DE      HLA-B*44/HLA-B*18 specific MHC-binding, tumour associated peptide #2.
XX
KW      tumour-associated peptide; Class I; major histocompatibility complex;
KW      MHC; HLA; cytostatic; vaccine; cytotoxic T lymphocyte; CTL; tumour;
KW      adenomatous disease; cancer; kidney; breast; pancreas; stomach; bladder;
KW      testis.
XX
OS      Homo sapiens.
XX
PN      WO2003102023-A1.
XX
PD      11-DEC-2003.
XX
PF      27-MAR-2003; 2003WO-EP003181.
XX
PR      29-MAY-2002; 2002DE-01025144.
XX
XX      (IMMA-) IMMATICS BIOTECHNOLOGIES GMBH.
XX
XX      Weinschenk T, Rammensee HG, Stevanovic S;
XX      WPI; 2004-053428/05.
XX
XX      New tumor-associated peptides, useful for diagnosis and immunotherapy of
XX      cancers, also nucleic acid encoding them.
XX
XX      Claim 1; SEQ ID NO 36; 78pp; German.
XX
XX      This invention describes novel tumour-associated peptides which can bind
XX      to a Class I molecule of the human major histocompatibility complex
XX      (MHC). MHC Class I-bound peptides were isolated from samples of kidney
XX      tumours from two patients, using HLA-specific antibodies, then sequenced
XX      by reverse-phase HPLC and mass spectrometry, to identify the new tumour-
```

CC associated peptides. The peptides have cytostatic activity and can be
CC used in vaccines or for the induction of cytotoxic T lymphocytes (CTL).
CC The peptides are used for treatment of tumors and/or adenomatous
CC diseases, especially cancer of kidney, breast, pancreas, stomach, bladder
CC and/or testis, for labelling leucocytes, especially T lymphocytes,
CC particularly for monitoring progress of tumour therapy, for preparation
CC of antibodies, also as diagnostic reagents for detecting specific
CC cytotoxic T lymphocytes in a cell population. Nucleic acids encoding the
CC tumour-associated peptides can also be used for genetic vaccination
CC against tumours. Cytotoxic T lymphocytes induced by the peptides in a
CC particular patient may kill cells of the same type in other patients or
CC cells of other tumour types. ADH48471-ADH48551 represent peptides
CC described in the disclosure of the invention.

XX SQ Sequence 9 AA;
Query Match 16.3%; Score 7; DB 8; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 EAAFLER 10
Db 2 EAAFLER 8
|||||

RESULT 8
ADI34277
ID ADI34277 standard; protein; 29 AA.
XX AC
XX ADI34277;
XX 15-APR-2004 (first entry)
XX DE Zipper peptide #2 for cross linking adenoviral ligands.
XX adenovirus vector; Cytostatic; HER2/neu; CD40; tumor necrosis factor;
XX TNF.
XX Unidentified.
XX WO2004009133-A1.
XX 29-JAN-2004.
XX 22-JUL-2003; 2003WO-US022852.
XX 22-JUL-2002; 2002US-0397951P.
XX (VECT-) VECTORLOGICS INC.
XX Korokhov N, Mikhoeva G.
XX WPI; 2004-132871/13.

XX Novel recombinant adenovirus having fiber protein modified by insertion
XX of first zipper peptide that can crosslink to second zipper peptide-
XX targeting ligand fusion protein, and binding between zipper peptides
XX targets vector to cell.
XX Claim 2; SEQ ID NO 2; 54pp; English.
XX The present invention relates to a targeted recombinant adenovirus
XX vector. The invention is useful for expressing a heterologous protein
XX chosen from a tumor associated antigen, HER2/neu and carcinoembryonic
XX antigen, in a target cell e.g., CD40 + cell such as dendritic cells, is
XX useful in a method of gene transfer to CD40 + cells, where the targeted
XX adenovirus vector mediates transfer of the gene encoding heterologous
XX protein to the cell such as the dendritic cell. The vector is useful in
XX gene therapy techniques for treatment of tumors. Multivalent interaction
XX or trimeric CD40L with CD40 receptors causes CD40 ligation, which then
XX results in enhanced survival of these cells and secretion of cytokines
XX such as interleukin (IL)-1, IL-6, IL-8, IL-10, IL-12, tumor necrosis
XX factor (TNF)- α agr, macrophage inflammatory protein (MIP)-1a and enzymes

CC such as matrix metalloproteinase. CD40-CD40L interaction also enhances
CC monocyte tumoricidal activity. The present sequence represents a zipper
CC peptide to be used for cross-linking targeting ligands with adenoviral
CC virions.
XX SQ Sequence 29 AA;
Query Match 16.3%; Score 7; DB 8; Length 29;
Best Local Similarity 100.0%; Pred. No. 5.8;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 TRVAELR 23
Db 14 TRVAELR 20
|||||

RESULT 9
AAG90757
ID AAG90757 standard; protein; 114 AA.
XX AC
XX AAG90757;
XX 26-SEP-2001 (first entry)
XX DE C glutamicum protein fragment SEQ ID NO: 4511.
XX Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
XX organic acid synthesis.
XX Corynebacterium glutamicum.
XX EP1108790-A2.
XX 20-JUN-2001.
XX 18-DEC-2000; 2000EP-00127688.
XX 16-DEC-1999; 99JP-00377484.
XX 07-APR-2000; 2000JP-00159162.
XX 03-AUG-2000; 2000JP-00280988.
XX (KYOW) KYOWA HAKKO KOGYO KK.
XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
XX Tateishi N, Senoh A, Ikeda M, Ozaki A;
XX WPI; 2001-376931/40.
XX N-PSDB; AAH65976.

XX Novel polynucleotides derived from Coryneform bacteria, for identifying
XX mutation point of a gene, measuring expression of a gene, analyzing
XX expression profile or pattern of a gene and identifying homologous gene.
XX Claim 17; SEQ ID NO 4511; 246pp + Sequence Listing; English.
XX The present invention provides a number of nucleotide and protein
XX sequences from the Coryneform bacterium Corynebacterium glutamicum. These
XX are useful for identifying the mutation point of a gene derived from a
XX mutant of coryneform bacterium, measuring expression amount and analysing
XX the expression profile or expression pattern of a gene derived from
XX Coryneform bacterium, and identifying a homologue of a gene derived from
XX coryneform bacterium. Coryneform bacteria are useful for producing amino
XX acids, nucleic acids, vitamins, saccharides and organic acids,
XX particularly L-lysine. The present sequence is a protein described in the
XX exemplification of the invention. Note: The sequence data for this patent
XX did not form part of the printed specification, but was obtained in
XX electronic format directly from the European Patent Office
XX Sequence 114 AA;
Query Match 16.3%; Score 7; DB 4; Length 114;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY      24 QRVQLR 30
DB      34 QRVQLR 40
      |||||
      |||||

RESULT 10
ADM18946
ID ADM18946 standard; protein; 186 AA.
XX
AC ADM18946;
XX
DT 01-JUL-2004 (first entry)
XX
DE Transposon Tn21 resolvase protein.
XX
KW gene therapy; serine recombinase; catalytic domain; DNA binding domain;
KW Tn3 resolvase; transgenic application.
XX
OS Transposon Tn21.
XX
PN WO2004029233-A2.
XX
PD 08-APR-2004.
XX
PF 25-SEP-2003; 2003WO-GB004169.
XX
PR 25-SEP-2002; 2002GB-00022229.
XX
PA (UNIU ) UNIV GLASGOW.
XX
PI Stark WM, Akopian A;
XX
DR WPI; 2004-316113/29.
XX
PT New serine recombinase having a catalytic domain mutated at Q105 and/or
PT G101 of Tn3 resolvase, and a DNA binding domain, for use in
PT biotechnology, gene therapy or transgenic applications.
XX
PS Disclosure; Fig 1; 95pp; English.
XX
CC The invention relates to a serine recombinase comprising a catalytic
CC domain and a DNA binding domain, where the catalytic domain is mutated at
CC G101 or at a position corresponding to G101 of Tn3 resolvase. The methods
CC and compositions of the present invention are useful in biotechnology,
CC gene therapy or transgenic applications. This sequence corresponds to a
CC transposase protein sequence used in the invention.
XX
SQ Sequence 186 AA;

      Query Match      16.3%; Score 7; DB 8; Length 186;
      Best Local Similarity 100.0%; Pred. No. 33;
      Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      20 AELRQRV 26
DB      152 AELRQRV 158
      |||||
      |||||

RESULT 11
ABO64927
ID ABO64927 standard; protein; 192 AA.
XX
AC ABO64927;
XX
DT 29-JUL-2004 (first entry)
XX
DE Klebsiella pneumoniae polypeptide seqid 11444.
XX
KW Recombinant expression vector; transcription regulatory element;
KW Klebsiella pneumoniae protein; antibacterial; vaccine.
XX
OS Klebsiella pneumoniae.

      Query Match      16.3%; Score 7; DB 8; Length 186;
      Best Local Similarity 100.0%; Pred. No. 33;
      Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      20 AELRQRV 26
DB      152 AELRQRV 158
      |||||
      |||||

RESULT 12
ABO82848
ID ABO82848 standard; protein; 217 AA.
XX
AC ABO82848;
XX
DT 29-JUL-2004 (first entry)
XX
DE Pseudomonas aeruginosa polypeptide #15023.
XX
KW Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
XX
OS Pseudomonas aeruginosa.
XX
PN US6551795-B1.
XX
PD 22-APR-2003.
XX
PF 18-FEB-1999; 99US-00252991.
XX
PR 18-FEB-1998; 98US-0074788P.
XX
PR 27-JUL-1998; 98US-0094190P.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
XX
PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;
XX
DR WPI; 2003-615309/58.
XX
DR N-PSDB; ABD16419.
XX
PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
PT useful as molecular targets for diagnostics, prophylaxis and treatment of
PT pathological conditions resulting from bacterial infection.
```

```
XX
PN US6610836-B1.
XX
PD 26-AUG-2003.
XX
PF 27-JAN-2000; 2000US-00489039.
XX
PR 29-JAN-1999; 99US-0117747P.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
XX
PI Breton GL, Osborne M;
XX
DR WPI; 2003-895346/82.
XX
DR N-PSDB; ACH98478.
XX
PT New nucleic acid encoding a Klebsiella pneumoniae polypeptide, useful for
PT preparing a vaccine composition against Klebsiella pneumoniae.
XX
PS Disclosure; SEQ ID NO 11444; 932pp; English.
XX
CC The invention describes a new isolated nucleic acid encoding a Klebsiella
CC pneumoniae polypeptide. Also described are: a recombinant expression
CC vector comprising the nucleic acid, operably linked to a transcription
CC regulatory element; and a cell comprising the recombinant expression
CC vector. The nucleic acid is useful for preparing a vaccine composition
CC against Klebsiella pneumoniae. This is the amino acid sequence of a
CC Klebsiella pneumoniae polypeptide of the invention
XX
SQ Sequence 192 AA;

      Query Match      16.3%; Score 7; DB 7; Length 192;
      Best Local Similarity 100.0%; Pred. No. 34;
      Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      20 AELRQRV 26
DB      158 AELRQRV 164
      |||||
      |||||

RESULT 12
ABO82848
ID ABO82848 standard; protein; 217 AA.
XX
AC ABO82848;
XX
DT 29-JUL-2004 (first entry)
XX
DE Pseudomonas aeruginosa polypeptide #15023.
XX
KW Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
XX
OS Pseudomonas aeruginosa.
XX
PN US6551795-B1.
XX
PD 22-APR-2003.
XX
PF 18-FEB-1999; 99US-00252991.
XX
PR 18-FEB-1998; 98US-0074788P.
XX
PR 27-JUL-1998; 98US-0094190P.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
XX
PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;
XX
DR WPI; 2003-615309/58.
XX
DR N-PSDB; ABD16419.
XX
PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
PT useful as molecular targets for diagnostics, prophylaxis and treatment of
PT pathological conditions resulting from bacterial infection.
```

XX Disclosure; SEQ ID NO 31594; 455pp; English.

PS The invention relates to Pseudomonas aeruginosa polypeptides and the

CC polynucleotides encoding them. The sequences are useful in diagnosis and

CC therapy of pathological conditions, as molecular targets for diagnostics,

CC prophylaxis and treatment of pathological conditions resulting from a

CC bacterial infection, for evaluating a compound, such as a polypeptide,

CC for the ability to bind a P. aeruginosa nucleic acid, as components of

CC effective antibacterial targets, as targets for antibacterial drugs,

CC including anti-P. aeruginosa drugs, as templates for recombinant

CC production of P. aeruginosa-derived peptides or polypeptides, as target

CC components for diagnosis and/or treatment of P. aeruginosa-caused

CC infection, and in detection of P. aeruginosa sequences or other sequences

CC of Pseudomonas species using biochip technology. Sequences ABO57826-

CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The

CC sequence data for this patent did not form part of the printed

CC specification but was obtained in electronic format from USPTO at

CC seqdata.uspto.gov/sequence.html

XX Sequence 217 AA;

SQ

Query Match 16.3%; Score 7; DB 7; Length 217;

Best Local Similarity 100.0%; Pred. No. 38;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 23 RQRVQL 29

DB 91 RQRVQL 97

|||||

RESULT 13

ADB74286

ID ADB74286 standard; protein; 341 AA.

XX

AC ADB74286;

XX

DT 04-DEC-2003 (first entry)

XX

DE Mycobacterium leprae non-naturally occurring peptide #9.

XX

KW Non-naturally occurring peptide; anion pump protein; tuberculosis;

KW hypersensitivity reaction; tuberculostatic.

XX

OS Mycobacterium leprae.

XX

PN US6583266-B1.

XX

PD 24-JUN-2003.

XX

PF 16-SEP-1994; 94US-00311731.

XX

PR 19-AUG-1993; 93US-00109181.

PR 22-OCT-1993; 93US-00142558.

XX

PA (GENO-) GENOME THERAPEUTICS CORP.

XX

PI Smith DR, Mao J;

XX

WPI; 2003-656441/62.

DR N-PSDB; ADB74274.

XX

PT New Mycobacterium tuberculosis anion pump peptide useful for as

PT tuberculosis vaccine and diagnosis of tuberculosis infection.

XX

PS Disclosure; SEQ ID NO 35; 26pp; English.

XX

CC The invention relates to a non-naturally occurring peptide of

CC Mycobacterium tuberculosis comprising an amino acid sequence

CC corresponding to an anion pump protein. The invention also relates to a

CC non-naturally occurring nucleic acid corresponding to a DNA sequence of

CC Mycobacterium tuberculosis or Mycobacterium leprae. The new peptide is

CC useful as a vaccine against Mycobacterium tuberculosis or Mycobacterium

CC

CC leprae or for screening for new tuberculosis drugs. Purified proteins

CC derived from the sequences of the invention may elicit a specific immune

CC response. The peptide may also be used to detect hypersensitivity

CC reactions of individuals exposed to Mycobacterium tuberculosis or

CC Mycobacterium leprae. The proteins and peptides may be affixed to solid

CC supports to detect antibodies typical of hypersensitivity reactions, from

CC a patient's sera. This sequence represents Mycobacterium leprae non-

CC naturally occurring peptide of the invention. Note: The sequence data for

CC this patent did not form part of the printed specification but was

CC obtained in electronic format directly from USPTO at

CC seqdata.uspto.gov/sequence.html.

XX Sequence 341 AA;

SQ

Query Match 16.3%; Score 7; DB 7; Length 341;

Best Local Similarity 100.0%; Pred. No. 59;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 IEAAFL 9

DB 95 IEAAFL 101

|||||

RESULT 14

ABB59427

ID ABB59427 standard; protein; 465 AA.

XX

AC ABB59427;

XX

DT 26-MAR-2002 (first entry)

XX

DE Drosophila melanogaster polypeptide SEQ ID NO 5073.

XX

KW Drosophila; developmental biology; cell signalling; insecticide;

KW pharmaceutical.

XX

OS Drosophila melanogaster.

XX

PN WO200171042-A2.

XX

PD 27-SEP-2001.

XX

PF 23-MAR-2001; 2001WO-US009231.

XX

PR 23-MAR-2000; 2000US-0191637P.

PR 11-JUL-2000; 2000US-00614150.

XX

PA (PEKE) PE CORP NY.

XX

PI Venter JC, Adams M, Li PWD, Myers EW;

XX

WPI; 2001-656860/75.

DR N-PSDB; ABL03530.

XX

PT New isolated nucleic acid detection reagent for detecting 1000 or more

PT genes from Drosophila and for elucidating cell signaling and cell-cell

PT interactions.

XX

PS Disclosure; SEQ ID NO 5073; 21pp + Sequence Listing; English.

XX

CC The invention relates to an isolated nucleic acid detection reagent

CC capable of detecting 1000 or more genes from Drosophila. The invention is

CC useful in developmental biology and in elucidating cell signalling and

CC cell-cell interactions in higher eukaryotes for the development of

CC insecticides, therapeutics and pharmaceutical drugs. The invention

CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA

CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB5773-

CC ABB72072). The sequence data for this patent did not form part of the

CC printed specification, but was obtained in electronic format directly

CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX

SQ Sequence 465 AA;

Search completed: July 26, 2005, 12:14:08
Job time : 111.694 secs

Query Match 16.3%; Score 7; DB 4; Length 465;
Best Local Similarity 100.0%; Pred. No. 79;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 5 AAFLERE 11
Db 167 AAFLERE 173

RESULT 15
ABO71770
ID ABO71770 standard; protein; 477 AA.
XX
AC ABO71770;
XX
DT 29-JUL-2004 (first entry)
XX
DE Pseudomonas aeruginosa polypeptide #3945.
XX
KW Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
XX
OS Pseudomonas aeruginosa.
XX
PN US6551795-B1.
XX
PD 22-APR-2003.
XX
PF 18-FEB-1999; 99US-00252991.
XX
PR 18-FEB-1998; 98US-0074788P.
XX
PR 27-JUL-1998; 98US-0094190P.
XX
FA (GENO-) GENOME THERAPEUTICS CORP.
XX
PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;
XX
XX WPI: 2003-615309/58.
XX
DR N-FSDB; ABD03341.
XX
PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
PT useful as molecular targets for diagnostics, prophylaxis and treatment of
PT pathological conditions resulting from bacterial infection.
XX
PS Disclosure; SEQ ID NO 20516; 455pp; English.
XX
XX The invention relates to Pseudomonas aeruginosa polypeptides and the
CC polynucleotides encoding them. The sequences are useful in diagnosis and
CC therapy of pathological conditions, as molecular targets for diagnostics,
CC prophylaxis and treatment of pathological conditions resulting from a
CC bacterial infection, for evaluating a compound, such as a polypeptide,
CC for the ability to bind a P. aeruginosa nucleic acid, as components of
CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-P. aeruginosa drugs, as templates for recombinant
CC production of P. aeruginosa-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of P. aeruginosa-caused
CC infection, and in detection of P. aeruginosa sequences or other sequences
CC of Pseudomonas species using biochip technology. Sequences ABO67826-
CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html
XX
SQ Sequence 477 AA;

Query Match 16.3%; Score 7; DB 7; Length 477;
Best Local Similarity 100.0%; Pred. No. 81;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 24 QRVQRLR 30
Db 197 QRVQRLR 203

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 26, 2005, 12:02:34 ; Search time 28.9653 Seconds
(without alignments)
110.819 Million cell updates/sec

Title: US-10-624-218-3
Perfect score: 43
Sequence: 1 LEIDAEFLERNTAETRA.....QRVQLRRNSQVTRYGPL 43

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 513545 seqs, 74649064 residues

Word size : 0

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pdp:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pdp:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pdp:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pdp:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pdp:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pdp:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	8	18.6	296	4	US-09-270-767-42478 Sequence 42478, A
2	7	16.3	192	4	US-09-489-039A-11444 Sequence 11444, A
3	7	16.3	217	4	US-09-252-991A-31594 Sequence 31594, A
4	7	16.3	323	4	US-09-270-767-42120 Sequence 42120, A
5	7	16.3	341	4	US-08-311-731A-35 Sequence 35, Appl
6	7	16.3	477	4	US-09-252-991A-20516 Sequence 20516, A
7	7	16.3	484	4	US-09-949-016-8999 Sequence 8999, Ap
8	7	16.3	485	4	US-09-252-991A-24868 Sequence 24868, A
9	7	16.3	515	3	US-08-942-012B-24 Sequence 24, Appl
10	7	16.3	532	1	US-08-285-440-5 Sequence 5, Appl
11	7	16.3	532	1	US-08-630-349-5 Sequence 5, Appl
12	7	16.3	538	4	US-09-949-016-6479 Sequence 6479, Ap
13	7	16.3	541	4	US-09-949-016-7217 Sequence 7217, Ap
14	7	16.3	558	1	US-08-285-440-6 Sequence 6, Appl
15	7	16.3	558	1	US-08-630-349-6 Sequence 6, Appl
16	7	16.3	793	4	US-09-538-092-1271 Sequence 1271, Ap
17	7	16.3	1069	4	US-09-902-540-11566 Sequence 11566, A
18	6	14.0	30	2	US-08-140-137A-41 Sequence 41, Appl
19	6	14.0	31	1	US-08-614-935-60 Sequence 60, Appl
20	6	14.0	31	3	US-09-130-287-60 Sequence 60, Appl
21	6	14.0	32	2	US-08-140-137A-27 Sequence 27, Appl
22	6	14.0	77	3	US-08-851-190-5 Sequence 5, Appl
23	6	14.0	77	4	US-09-403-861A-12 Sequence 12, Appl
24	6	14.0	92	4	US-09-198-452A-1170 Sequence 1170, Ap
25	6	14.0	97	4	US-09-690-454-211 Sequence 211, App
26	6	14.0	115	3	US-09-247-155-124 Sequence 124, App
27	6	14.0	115	4	US-09-471-276-1558 Sequence 1558, Ap

28	6	14.0	119	1	US-08-702-344-2 Sequence 2, Appl
29	6	14.0	122	4	US-09-949-016-7298 Sequence 5298, Ap
30	6	14.0	133	3	US-08-966-318-5 Sequence 5, Appl
31	6	14.0	133	3	US-09-216-619-5 Sequence 5, Appl
32	6	14.0	134	3	US-08-851-190-1 Sequence 1, Appl
33	6	14.0	134	4	US-09-403-861A-6 Sequence 6, Appl
34	6	14.0	136	4	US-09-902-540-10119 Sequence 10119, A
35	6	14.0	137	4	US-09-403-861A-2 Sequence 2, Appl
36	6	14.0	153	4	US-09-248-796A-18633 Sequence 18633, A
37	6	14.0	154	4	US-09-902-540-12708 Sequence 12708, A
38	6	14.0	159	4	US-09-270-767-40847 Sequence 40847, A
39	6	14.0	159	4	US-09-270-767-56063 Sequence 56063, A
40	6	14.0	186	4	US-09-902-540-13724 Sequence 13724, A
41	6	14.0	193	4	US-09-438-185A-222 Sequence 222, App
42	6	14.0	201	4	US-09-252-991A-24377 Sequence 24377, A
43	6	14.0	208	4	US-09-270-767-56953 Sequence 56953, A
44	6	14.0	209	3	US-09-134-001C-5126 Sequence 5126, Ap
45	6	14.0	213	4	US-09-902-540-15826 Sequence 15826, A

ALIGNMENTS

RESULT 1
US-09-270-767-42478
; Sequence 42478, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 42478
; LENGTH: 296
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-42478

Query Match 18.6%; Score 8; DB 4; Length 296;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	15	LET RVAEL 22
DB	91	LET RVAEL 98

RESULT 2
US-09-489-039A-11444
; Sequence 11444, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR FILING DATE: 2000-01-27
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 11444
; LENGTH: 192
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-11444

Query Match 16.3%; Score 7; DB 4; Length 192;

```
Best Local Similarity 100.0%; Pred. No. 9.2;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 AELRQV 26
Db 158 AELRQV 164
|||||

RESULT 3
US-09-252-991A-31594
; Sequence 31594, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 31594
; LENGTH: 217
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31594

Query Match 16.3%; Score 7; DB 4; Length 217;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 23 RQRVQL 29
Db 91 RQRVQL 97
|||||

RESULT 4
US-09-270-767-42120
; Sequence 42120, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 42120
; LENGTH: 323
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-42120

Query Match 16.3%; Score 7; DB 4; Length 323;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 AAFLE 11
Db 182 AAFLE 188
|||||

RESULT 5
US-08-311-731A-35
; Sequence 35, Application US/08311731A
; Patent No. 6583266
; GENERAL INFORMATION:
; APPLICANT: SMITH, DOUGLAS
; APPLICANT: MAO, JEN-I
```

```
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; TITLE OF INVENTION: RELATING TO MYCOBACTERIUM TUBERCULOSIS AND LAPRAE FOR
; NUMBER OF SEQUENCES: 411
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: WOLF, GREENFIELD & SACKS, P.C.
; STREET: 600 ATLANTIC AVENUE
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 02210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/311,731A
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: GATES, EDWARD R.
; REGISTRATION NUMBER: 31,616
; REFERENCE/DOCKET NUMBER: C0044/7125
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/720-3500
; TELEFAX: 617/720-2441
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 341 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: MYCOBACTERIUM LEPRAE
; US-08-311-731A-35

Query Match 16.3%; Score 7; DB 4; Length 341;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 IEAAFL 9
Db 95 IEAAFL 101
|||||

RESULT 6
US-09-252-991A-20516
; Sequence 20516, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 20516
; LENGTH: 477
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-20516

Query Match 16.3%; Score 7; DB 4; Length 477;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Oy 24 QRVORLR 30
Db 197 QRVORLR 203

RESULT 7

US-09-949-016-8999
; Sequence 8999, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8999
; LENGTH: 484
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-8999

Query Match 16.3%; Score 7; DB 4; Length 484;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 4 EAAFLER 10
Db 96 EAAFLER 102

RESULT 8

US-09-252-991A-24868
; Sequence 24868, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 24868
; LENGTH: 485
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24868

Query Match 16.3%; Score 7; DB 4; Length 485;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2 EIEAAFL 8
Db 428 EIEAAFL 434

RESULT 9

US-08-942-012B-24
; Sequence 24, Application US/08942012B

Patent No. 6235278
; GENERAL INFORMATION:
; APPLICANT: Miller, Lois K.
; APPLICANT: Lu, Albert
; APPLICANT: Dierks, Peter
; APPLICANT: Black, Bruce
; TITLE OF INVENTION: Biological Insect Control Agents Expressing
; Insect-Specific Toxin Genes, Methods and Compositions
; FILE REFERENCE: 28-96a
; CURRENT APPLICATION NUMBER: US/08/942,012B
; CURRENT FILING DATE: 1997-10-01
; PRIOR APPLICATION NUMBER: 08/729,606
; PRIOR FILING DATE: 2000-10-01
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 24
; LENGTH: 515
; TYPE: PRT
; ORGANISM: Helicoverpa zea nuclear polyhedrosis virus
US-08-942-012B-24

Query Match 16.3%; Score 7; DB 3; Length 515;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 21 ELRQRVQ 27
Db 237 ELRQRVQ 243

RESULT 10

US-08-285-440-5
; Sequence 5, Application US/08285440
; Patent No. 5532337
; GENERAL INFORMATION:
; APPLICANT: Ken'ichiro HAYASHI et al.
; TITLE OF INVENTION: POLYPEPTIDE
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 500 kb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: DisplayWrite
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/285,440
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/858,947
; FILING DATE: March 27, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek, Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX:
; TELEX:
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 532 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL:

;; PRIOR FILING DATE: 2000-10-03
;; PRIOR APPLICATION NUMBER: 60/231,498
;; PRIOR FILING DATE: 2000-09-08
;; NUMBER OF SEQ ID NOS: 207012
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 7217
;; LENGTH: 541
;; TYPE: PRT
;; ORGANISM: Human
US-09-949-016-7217

Query Match 16.3%; Score 7; DB 4; Length 541;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 EAAFLER 10
Db 96 EAAFLER 102

RESULT 14
US-08-285-440-6
; Sequence 6, Application US/08285440
; Patent No. 5532337
; GENERAL INFORMATION:
; APPLICANT: Ken'ichiro HAYASHI et al.
; TITLE OF INVENTION: POLYPEPTIDE
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005

COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 500 kb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: DisplayWrite
; CURRENT APPLICATION DATA:
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/858,947
; FILING DATE: March 27, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek, Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX:
; TELEX:

INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 558 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHEetical:
; ANTI-SENSE:
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
; ORGANISM:
; STRAIN:
; INDIVIDUAL ISOLATE:
; DEVELOPMENTAL STAGE:
; HAPLOTYPE:
; TISSUE TYPE:
; CELL TYPE:

;; CELL LINE:
;; ORGANELLE:
;; IMMEDIATE SOURCE:
;; LIBRARY:
;; CLONE:
;; POSITION IN GENOME:
;; CHROMOSOME/SEGMENT:
;; MAP POSITION:
;; UNITS:
;; FEATURE:
;; NAME/KEY:
;; LOCATION:
;; IDENTIFICATION METHOD:
;; OTHER INFORMATION:
;; PUBLICATION INFORMATION:
;; AUTHORS:
;; TITLE:
;; JOURNAL:
;; VOLUME:
;; ISSUR:
;; PAGES:
;; DATE:
;; DOCUMENT NUMBER:
;; FILING DATE:
;; PUBLICATION DATE:
;; RELEVANT RESIDUES IN SEQ ID NO:
US-08-285-440-6

Query Match 16.3%; Score 7; DB 1; Length 558;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 EAAFLER 10
Db 87 EAAFLER 93

RESULT 15
US-08-630-349-6
; Sequence 6, Application US/08630349
; Patent No. 5739008
; GENERAL INFORMATION:
; APPLICANT: Ken'ichiro HAYASHI et al.
; TITLE OF INVENTION: POLYPEPTIDE
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 144 mb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/630,349
; FILING DATE: April 10, 1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/285,440
; FILING DATE: August 4, 1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/858,947
; FILING DATE: March 27, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek, Jr.
; REGISTRATION NUMBER: 33,367
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 558 amino acids

Thu Jul 28 11:34:43 2005

```
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-630-349-6
Query Match      16.3%; Score 7; DB 1; Length 558;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy      4 EAAFLR 10
Db      87 EAAFLR 93
```

Search completed: July 26, 2005, 12:22:31
Job time : 29.9653 secs

RESULT 2

```
US-10-437-963-204275
; Sequence 204275, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 204275
; LENGTH: 249
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(249)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_99379C.1.pep
US-10-437-963-204275

Query Match      18.6%; Score 8; DB 16; Length 249;
Best Local Similarity 100.0%; Pred. No. 6.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      25 RVQRLNR 32
      |||||
Db      148 RVQRLNR 155

RESULT 3
US-10-425-114-72646
; Sequence 72646, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 72646
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3595-019-D9_FLI.pep
US-10-425-114-72646

Query Match      18.6%; Score 8; DB 15; Length 432;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      22 LRQRVQRL 29
      |||||
Db      75 LRQRVQRL 82
```

```
RESULT 4
US-10-437-963-109382
; Sequence 109382, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 109382
; LENGTH: 623
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(623)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_13545C.1.pep
US-10-437-963-109382

Query Match      18.6%; Score 8; DB 16; Length 623;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      25 RVQRLNR 32
      |||||
Db      510 RVQRLNR 517

RESULT 5
US-10-437-963-168645
; Sequence 168645, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 168645
; LENGTH: 644
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(644)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_6713C.1.pep
US-10-437-963-168645

Query Match      18.6%; Score 8; DB 16; Length 644;
Best Local Similarity 100.0%; Pred. No. 16;
```


Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 RVQRLNR 32
|||||||
Db 34 RVQRLNR 41

RESULT 6

US-10-437-963-173197
; Sequence 173197, Application US/10437963
; Publication No. US20040123343A1

; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; APPLICANT: Wu, Wei

; APPLICANT: Boukharov, Andrey A.

; APPLICANT: Barbazuk, Brad

; APPLICANT: Li, Ping

; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B

; CURRENT APPLICATION NUMBER: US/10/437,963

; CURRENT FILING DATE: 2003-05-14

; NUMBER OF SEQ ID NOS: 204966

; SEQ ID NO 173197

; LENGTH: 653

; TYPE: PRT

; ORGANISM: Oryza sativa

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT_MRT4530_71259C.1.pep

US-10-437-963-173197

Query Match 18.6%; Score 8; DB 16; Length 653;

Best Local Similarity 100.0%; Pred. No. 17;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 RVQRLNR 32
|||||||
Db 461 RVQRLNR 469

RESULT 7

US-10-437-963-103617

; Sequence 103617, Application US/10437963

; Publication No. US20040123343A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; APPLICANT: Wu, Wei

; APPLICANT: Boukharov, Andrey A.

; APPLICANT: Barbazuk, Brad

; APPLICANT: Li, Ping

; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B

; CURRENT APPLICATION NUMBER: US/10/437,963

; CURRENT FILING DATE: 2003-05-14

; NUMBER OF SEQ ID NOS: 204966

; SEQ ID NO 103617

; LENGTH: 700

; TYPE: PRT

; ORGANISM: Oryza sativa

; FEATURE:

; NAME/KEY: unsure

; LOCATION: (1)..(700)

; OTHER INFORMATION: unsure at all xaa locations

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT_MRT4530_101028C.1.pep

US-10-437-963-103617

Query Match 18.6%; Score 8; DB 16; Length 700;

Best Local Similarity 100.0%; Pred. No. 18;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 RVQRLNR 32
|||||||
Db 334 RVQRLNR 341

RESULT 8

US-10-437-963-204238

; Sequence 204238, Application US/10437963

; Publication No. US20040123343A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; APPLICANT: Wu, Wei

; APPLICANT: Boukharov, Andrey A.

; APPLICANT: Barbazuk, Brad

; APPLICANT: Li, Ping

; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B

; CURRENT APPLICATION NUMBER: US/10/437,963

; CURRENT FILING DATE: 2003-05-14

; NUMBER OF SEQ ID NOS: 204966

; SEQ ID NO 204238

; LENGTH: 901

; TYPE: PRT

; ORGANISM: Oryza sativa

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT_MRT4530_99345C.1.pep

US-10-437-963-204238

Query Match 18.6%; Score 8; DB 16; Length 901;

Best Local Similarity 100.0%; Pred. No. 22;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 RVQRLNR 32
|||||||
Db 23 RVQRLNR 30

RESULT 9

US-10-437-963-103572

; Sequence 103572, Application US/10437963

; Publication No. US20040123343A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; APPLICANT: Wu, Wei

; APPLICANT: Boukharov, Andrey A.

; APPLICANT: Barbazuk, Brad

; APPLICANT: Li, Ping

; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B

; CURRENT APPLICATION NUMBER: US/10/437,963

; CURRENT FILING DATE: 2003-05-14

; NUMBER OF SEQ ID NOS: 204966

; SEQ ID NO 103572

; LENGTH: 959

; TYPE: PRT

; ORGANISM: Oryza sativa

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT_MRT4530_100989C.1.pep

US-10-437-963-103572

Query Match 18.6%; Score 8; DB 16; Length 959;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 RVQRLNR 32
Db 420 RVQRLNR 427

RESULT 10
US-10-437-963-156507
; Sequence 156507, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 156507
; TYPE: PRT
; ORGANISM: Oryza sativa
; NAME/KEY: unsure
; LOCATION: (1)..(1029)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_56168C.1.pep
US-10-437-963-156507

Query Match 18.6%; Score 8; DB 16; Length 1029;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 RVQRLNR 32
Db 29 RVQRLNR 36

RESULT 11
US-10-437-963-204316
; Sequence 204316, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 204316
; LENGTH: 1032
; TYPE: PRT
; ORGANISM: Oryza sativa
; NAME/KEY: unsure
; LOCATION: (1)..(1029)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_56168C.1.pep
US-10-437-963-156507

; OTHER INFORMATION: Clone ID: PAT_MRT4530_99415C.1.pep
US-10-437-963-204316

Query Match 18.6%; Score 8; DB 16; Length 1032;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 RVQRLNR 32
Db 23 RVQRLNR 30

RESULT 12
US-10-437-963-204228
; Sequence 204228, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 204228
; LENGTH: 1074
; TYPE: PRT
; ORGANISM: Oryza sativa
; NAME/KEY: unsure
; LOCATION: (1)..(1029)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_99336C.1.pep
US-10-437-963-204228

Query Match 18.6%; Score 8; DB 16; Length 1074;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 RVQRLNR 32
Db 224 RVQRLNR 231

RESULT 13
US-10-437-963-185083
; Sequence 185083, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 185083
; LENGTH: 1094
; TYPE: PRT
; ORGANISM: Oryza sativa
; NAME/KEY: unsure
; LOCATION: (1)..(1029)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_82013C.1.pep
US-10-437-963-185083

US-10-437-963-185083

Query Match 18.6%; Score 8; DB 16; Length 1094;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 25 RVQRLNR 32
Db 373 RVQRLNR 380

RESULT 14

US-09-815-242-11828
US-09-815-242-11828, Application US/09815242
; Sequence 11828, Application US/09815242
; Patent No. US20020061569A1

GENERAL INFORMATION:

APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard

TITLE OF INVENTION: Identification of Essential Genes in Prokaryotes

FILE REFERENCE: ELITRA.011A

CURRENT APPLICATION NUMBER: US/09/815,242

CURRENT FILING DATE: 2001-03-21

PRIOR APPLICATION NUMBER: 60/191,078

PRIOR FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: 60/206,848

PRIOR FILING DATE: 2000-05-23

PRIOR APPLICATION NUMBER: 60/207,727

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/242,578

PRIOR FILING DATE: 2000-10-23

PRIOR APPLICATION NUMBER: 60/253,625

PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/257,931

PRIOR FILING DATE: 2000-12-22

PRIOR APPLICATION NUMBER: 60/269,308

PRIOR FILING DATE: 2001-02-16

NUMBER OF SEQ ID NOS: 14110

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 11828

LENGTH: 1162

TYPE: PRT

ORGANISM: Pseudomonas aeruginosa

US-09-815-242-11828

Query Match 18.6%; Score 8; DB 9; Length 1162;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 14 ALETRVAE 21
Db 665 ALETRVAE 672

RESULT 15

US-10-282-122A-66294

Sequence 66294, Application US/10282122A

Publication No. US20040029129A1

GENERAL INFORMATION:

APPLICANT: Wang, Liangsu

APPLICANT: Zamudio, Carlos

APPLICANT: Malone, Cheryl

APPLICANT: Haselbeck, Robert

APPLICANT: Ohlsen, Kari

APPLICANT: Zyskind, Judith

APPLICANT: Wall, Daniel

APPLICANT: Trawick, John

APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 66294
LENGTH: 1162
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-10-282-122A-66294

Query Match 18.6%; Score 8; DB 15; Length 1162;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 14 ALETRVAE 21
Db 665 ALETRVAE 672

Search completed: July 26, 2005, 12:28:16
Job time : 102.229 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 26, 2005, 12:00:54 ; Search time 19.7083 Seconds
(without alignments)
209.927 Million cell updates/sec

Title: US-10-624-218-3
Perfect score: 43
Sequence: 1 LEIBAAFLERNTALETRVA.....QRVORLRNRVSQYTRYGPL 43

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 283416 seqs, 96216763 residues

Word size : 0
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : PIR_79:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	8	18.6	1162	2 D83454	conserved hypothe
2	7	16.3	186	1 RP8C21	resolvase - Escher
3	7	16.3	340	2 H07090	hypotheical prote
4	7	16.3	341	2 E87197	probable anion tra
5	7	16.3	362	2 E75412	spermidine/putresc
6	7	16.3	364	2 E75532	periplasmic serine
7	7	16.3	514	2 T44976	hydantoinease homol
8	7	16.3	790	1 G69071	heavy-metal-transp
9	7	16.3	793	1 JH0628	caldesmon - human
10	7	16.3	1155	2 G87477	transcription-repa
11	7	16.3	2712	2 T05113	hypotheical prote
12	6	14.0	67	2 F83430	type III export pr
13	6	14.0	74	2 AE1003	hypotheical prote
14	6	14.0	74	2 AE1953	SLYX protein homol
15	6	14.0	74	2 CB1010	slyx protein, prob
16	6	14.0	77	2 S36827	DSIP-immunoreactiv
17	6	14.0	94	2 G75493	hypotheical prote
18	6	14.0	95	2 F87302	CheX protein [impo
19	6	14.0	96	2 AE3640	hypotheical prote
20	6	14.0	99	2 B46598	ski-related protei
21	6	14.0	108	2 A13599	hypotheical prote
22	6	14.0	126	2 E70420	NADH2 dehydrogenas
23	6	14.0	130	2 T14749	hypotheical prote
24	6	14.0	133	2 F88492	protein T07E3.7 (i
25	6	14.0	138	2 A12996	lactoylglutathione
26	6	14.0	141	2 T34865	probable oxidoredu
27	6	14.0	141	2 AH2725	hypotheical prote
28	6	14.0	142	2 JQ1031	insulin receptor -
29	6	14.0	145	2 H98286	glyoxalase I relat

30	6	14.0	153	2 T40951	40s ribosomal prot
31	6	14.0	154	2 T37489	40s ribosomal prot
32	6	14.0	157	2 B71966	hypotheical prote
33	6	14.0	157	2 G64540	hypotheical prote
34	6	14.0	163	2 AH3002	shikimate kinase (
35	6	14.0	166	2 H98280	shikimate kinase (
36	6	14.0	168	2 T50922	bZIP protein HY5 (
37	6	14.0	169	2 T30684	probable dual spec
38	6	14.0	175	2 E97553	hypotheical prote
39	6	14.0	177	2 C87468	response regulator
40	6	14.0	180	2 I48129	xe169 (escapes X-i
41	6	14.0	180	2 I84689	escapes X-inactiva
42	6	14.0	181	2 D86518	hypotheical prote
43	6	14.0	181	2 C72104	hypotheical prote
44	6	14.0	185	2 T45083	pyruvate synthase
45	6	14.0	185	2 G71113	probable pyruvate

ALIGNMENTS

RESULT 1

D83454
conserved hypotheical protein PA1527 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C:Accession: D83454
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: D83454
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1162 <STO>
A:Cross-references: UNIPROT:Q91316; GB:AE004581; GB:AE004091; NID:g9947482; PIDN:AAG0491
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA1527

Query Match 18.6%; Score 8; DB 2; Length 1162;
Best Local Similarity 100.0%; Pred. No. 3.8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	14	ALETRVAE	21
DB	665	ALETRVAE	672

RESULT 2

RP8C21
resolvase - Escherichia coli transposons
C:Species: Escherichia coli
C:Date: 04-Dec-1986 #sequence_revision 04-Dec-1986 #text_change 09-Jul-2004
C:Accession: A03543; S09631
R:Hyde, D.R.; Tu, C.P.T.
Cell 42, 629-638, 1985
A:Title: tnpM: a novel regulatory gene that enhances Tn21 transposition and suppresses
A:Reference number: A90871; MUID:85282620; PMID:2992807
A:Accession: A03543
A:Molecule type: DNA
A:Residues: 1-186 <HYD>
A:Cross-references: UNIPROT:P04130; GB:AF071413; NID:g3513654; PIDN:AAC33925.1; PID:g351
A:Experimental source: transposon Tn21
R:Diver, W.P.; Grinstead, J.; Fritzinger, D.C.; Brown, N.L.; Altenbuchner, J.; Rogowsky,
Mol. Gen. Genet. 191, 189-193, 1983
A:Title: DNA sequences of and complementation by the tnpR genes of Tn21, Tn501 and Tn17;
A:Reference number: S07292; MUID:84013495; PMID:6312271
A:Accession: S09631
A>Status: preliminary
A:Molecule type: DNA

A;Residues: 1-186 <DIV>
A;Cross-references: EMBL:X01298; NID:943707; PIDN:CAA25626.1; PID:943708
A;Experimental source: transposon Tn1721
C;Genetics:
A;Gene: tpmR
A;Mobile element: transposons
C;Function:
A;Description: protein catalyzes the resolution (a site-specific recombination) of the d
mediate of the transposition process of Tn21, a transposon encoding resistance to sulfa
C;Superfamily: transposase repressor
C;Keywords: DNA binding; site-specific integration; transposition

Query Match 16.3%; Score 7; DB 1; Length 186;
Best Local Similarity 100.0%; Pred. No. 7.9;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 20 AELRQV 26
DB 152 AELRQV 158
|||||

RESULT 3
H70790
hypothetical protein Rv3679 - Mycobacterium tuberculosis (strain H37RV)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C;Accession: H70790
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.
Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holtroyd, S.
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A;Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Accession: H70790
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-340 <COL>
A;Cross-references: UNIPROT:O69647; GB:AL022121; GB:AL123456; NID:93261559; PIDN:CAA1800
A;Experimental source: strain H37RV
C;Genetics:
A;Gene: Rv3679

Query Match 16.3%; Score 7; DB 2; Length 340;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 IEAAFL 9
DB 95 IEAAFL 101
|||||

RESULT 4
E87197
probable anion transporter protein ML2305 [imported] - Mycobacterium leprae
C;Species: Mycobacterium leprae
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C;Accession: E87197
R;Cole, S.T.; Eiglmeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; Ho
R.; Davies, R.M.; Devlin, K.; Duthoy, S.; Feltwell, T.; Fraser, A.; Hamlin, N.; Holtroyd,
eam, M.A.; Rutherford, K.M.
Nature 409, 1007-1011, 2001
A;Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; Sq
A;Title: Massive gene decay in the leprosy bacillus
A;Reference number: A86909; MUID:21128732; PMID:11234002
A;Accession: E87197
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-341 <STO>
A;Cross-references: UNIPROT:Q9CB88; GB:AL450380; NID:913093930; PIDN:CAC31821.1; GSPDB:C
C;Genetics:
A;Gene: ML2305

Query Match 16.3%; Score 7; DB 2; Length 341;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 IEAAFL 9
DB 95 IEAAFL 101
|||||

RESULT 5

E75412
spermidine/putrescine ABC transporter, ATP-binding protein - Deinococcus radiodurans (str
C;Species: Deinococcus radiodurans
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 16-Aug-2004
C;Accession: E75412
R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; I
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A;Reference number: A75250; MUID:20036896; PMID:10567266
A;Accession: E75412
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-362 <WHI>
A;Cross-references: UNIPROT:Q9RUT0; GB:AE001977; GB:AE000513; NID:96459045; PIDN:AAF1087;
A;Experimental source: strain R1
C;Genetics:
A;Gene: DRI302
A;Map position: 1
C;Superfamily: ATP-binding cassette homology

Query Match 16.3%; Score 7; DB 2; Length 362;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 AELRQV 26
DB 110 AELRQV 116
|||||

RESULT 6

D75532
periplasmic serine proteinase, HtrA/DegQ/DegS family - Deinococcus radiodurans (strain R1
C;Species: Deinococcus radiodurans
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: D75532
R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; I
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A;Reference number: A75250; MUID:20036896; PMID:10567266
A;Accession: D75532
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-364 <WHI>
A;Cross-references: UNIPROT:Q9RX16; GB:AE001893; GB:AE000513; NID:96458001; PIDN:AAF09090;
A;Experimental source: strain R1
C;Genetics:
A;Gene: DR0327
A;Map position: 1
C;Superfamily: Escherichia coli trypsin-like proteinase degS; GLGF domain homology; trypt

Query Match 16.3%; Score 7; DB 2; Length 364;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 TRVAEL 23
DB 234 TRVAEL 240
|||||

RESULT 7

T44976
 hydantoinase homolog [imported] - Haloferax volcanii megaplasmid pHV3
 C;Species: Haloferax volcanii
 C;Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 09-Jul-2004
 C;Accession: T44976
 R;Parahani, R.; Imbeault, J.C.; St Jean, A.; Chan, C.C.Y.; Allard, G.; Charlebois, R.L.
 submitted to the EMBL Data Library, March 1997
 A;Description: Hereditary instability of the megaplasmid pHV3, and filamentation in the
 A;Reference number: 222886
 A;Accession: T44976
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-514 <FAR>
 A;Cross-references: UNIPROT:O34180; EMBL:U95372; PIDN:AAB71802.1
 A;Experimental source: strain DS2
 C;Genetics:
 A;Map position: megaplasmid pHV3
 A;Genome: plasmid
 A;Note: expressed during exponential growth

Query Match 16.3%; Score 7; DB 2; Length 514;
 Best Local Similarity 100.0%; Pred. No. 20;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 ETRVAEL 22

Db 163 ETRVAEL 169

RESULT 8

G69071
 heavy-metal-transporting ATPase (EC 3.6.1.-) MTH1535 - Methanobacterium thermoautotroph
 C;Species: Methanobacterium thermoautotrophicum
 C;Date: 16-Oct-1998 #sequence_revision 16-Oct-1998 #text_change 09-Jul-2004
 C;Accession: G69071
 R;Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;
 Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.
 ki, S.; Church, G.M.; Daniel, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
 J. Bacteriol. 179, 7135-7155, 1997
 A;Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: func
 A;Reference number: A69000; MUID:98037514; PMID:9371463
 A;Accession: G69071
 A;Status: nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-790 <MTH>
 A;Cross-references: UNIPROT:O27578; GB:AE000913; GB:AE000666; NID:92622646; PIDN:AAB8600
 A;Experimental source: strain Delta H
 C;Genetics:
 A;Gene: MTH1535
 C;Function:
 A;Description: transition metal ion transport coupled with ATP hydrolysis
 C;Superfamily: Bacillus probable copper-transporting ATPase yvgX; ATPase nucleotide-bind
 C;Keywords: ATP; copper transport; hydrolase; ion transport; phosphoprotein; transmembra
 P;8-37/Domain: heavy-metal-associated homology <HMA1>
 P;76-105/Domain: heavy-metal-associated homology <HMA2>
 P;183-515/Domain: ATPase transduction domain homology <ATT>
 P;590-732/Domain: ATPase nucleotide-binding domain homology <ATN>
 P;13.16/Binding site: transition metal ions (Cys) #status predicted
 P;81.84/Binding site: transition metal ions (Cys) #status predicted
 P;331/Active site: Glu #status predicted
 P;477/Active site: Asp (aspartylphosphate intermediate) #status predicted

Query Match 16.3%; Score 7; DB 1; Length 790;
 Best Local Similarity 100.0%; Pred. No. 29;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 TRVAELR 23

Db 650 TRVAELR 656

RESULT 9

JH0628

caldesmon - human
 N;Alternate names: nonmuscle caldesmon; smooth muscle caldesmon
 N;Contains: h-caldesmon; l-caldesmon
 C;Species: Homo sapiens (man)
 C;Date: 17-Aug-1992 #sequence_revision 27-Jun-1994 #text_change 09-Jul-2004
 C;Accession: JH0628; A41186; B47193; A47193
 R;Humphrey, M.B.; Herrera-Sosa, H.; Gonzalez, G.; Lee, R.; Bryan, J.
 Gene 112, 197-204, 1992
 A;Title: Cloning of cDNAs encoding human caldesmons.
 A;Reference number: JH0628; MUID:92209999; PMID:1555769
 A;Accession: JH0628
 A;Molecule type: mRNA
 A;Residues: 1-793 <HUM>
 A;Cross-references: UNIPROT:Q05682; GB:M83216; NID:9306508; PIDN:AAA58419.1; PID:gl80195
 A;Experimental source: aorta
 R;Novy, R.E.; Lin, J.L.C.; Lin, J.J.C.
 J. Biol. Chem. 266, 16917-16924, 1991
 A;Title: Characterization of cDNA clones encoding a human fibroblast caldesmon isoform a
 A;Reference number: A41186; MUID:91358497; PMID:1885618
 A;Accession: A41186
 A;Molecule type: mRNA
 A;Residues: 1-207,463-529, 'M', 531-793 <NOV>
 A;Cross-references: GB:M64110; NID:gl79829; PIDN:AAA35636.1; PID:gl79830
 A;Experimental source: fetal lung
 R;Hayashi, K.; Yano, H.; Hashida, T.; Takeuchi, R.; Takeda, O.; Asada, K.; Takahashi, E.
 Proc. Natl. Acad. Sci. U.S.A. 89, 12122-12126, 1992
 A;Title: Genomic structure of the human caldesmon gene.
 A;Reference number: A47193; MUID:93101679; PMID:1465449
 A;Accession: B47193
 A;Molecule type: DNA
 A;Residues: 74-436 <HA2>
 A;Experimental source: placenta
 A;Note: sequence extracted from NCBI backbone (NCBI:120706, NCBIP:120707)
 C;Comment: Because of its ability to inhibit the ATPase activity of actomyosin, caldesmo
 C;Comment: The binding of caldesmon to F-actin is modulated by calcium and calmodulin.
 C;Comment: Two calmodulin molecules can bind to nonoverlapping domains of each caldesmon
 C;Genetics:
 A;Gene: GDB:CALD1
 A;Cross-references: GDB:133717; OMIM:114213
 A;Map position: 7q33-7q33
 C;Superfamily: caldesmon
 C;Keywords: actin binding; alternative splicing; calmodulin binding; muscle; phosphopro
 P;1-793/Product: h-caldesmon #status predicted <MAT>
 P;1-207,463-793/Product: l-caldesmon #status predicted <MA2>
 P;257-389/Region: 13-residue repeats
 P;638,730,753/Binding site: phosphate (Thr) (covalent) #status predicted
 P;724,759/Binding site: phosphate (Ser) (covalent) #status predicted
 Query Match 16.3%; Score 7; DB 1; Length 793;
 Best Local Similarity 100.0%; Pred. No. 29;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 EAAFLER 10

Db 93 EAAFLER 99

RESULT 10

G87477
 transcription-repair coupling factor [imported] - Caulobacter crescentus
 C;Species: Caulobacter crescentus
 C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
 C;Accession: G87477
 R;Niernman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
 B.; Iand, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolo
 n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
 Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
 A;Title: Complete Genome Sequence of Caulobacter crescentus.
 A;Reference number: A87249; MUID:21173698; PMID:11259647
 A;Accession: G87477
 A;Status: preliminary
 A;Molecule type: DNA

A;Residues: 1-1155 <STO>
A;Cross-references: UNIPROT:Q9A782; GB:AE005673; NID:gl3423283; PIDN:AAK23819.1; GSPDB:C
C;Genetics:
A;Gene: C01844
C;Superfamily: transcription-repair coupling protein

Query Match 16.3%; Score 7; DB 2; Length 1155;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 VAEQR 25
|||||
Db 989 VAEQR 995

RESULT 11
T05113
hypothetical protein F28M20.240 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C;Accession: T05113
R;Bevan, M.; Rieger, M.; Mueller-Auer, S.; Zipp, M.; Schaefer, M.; Hoheisel, J.; Mewes, S.; Moller, S.; O'Gaora, P.
A;Reference number: 215398
A;Accession: T05113
A;Molecule type: DNA
A;Residues: 1-2712 <BEV>
A;Cross-references: UNIPROT:Q9SB74; EMBL:AL031004
A;Experimental source: cultivar Columbia; BAC clone F28M20
C;Genetics:
A;Map position: 4
A;Introns: 17/3; 240/1; 1950/1; 2118/3; 2381/1; 2599/3; 2645/1; 2679/2
A;Note: F28M20.240

Query Match 16.3%; Score 7; DB 2; Length 2712;
Best Local Similarity 100.0%; Pred. No. 89;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 RVAEQR 24
|||||
Db 2561 RVAEQR 2567

RESULT 12
F83430
type III export protein PscE PA1718 [imported] - Pseudomonas aeruginosa (strain PAO1)
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C;Accession: F83430
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Boman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, J.; Lory, S.; Olson, M.V.
A;Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic pathogen
A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: F83430
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-67 <STO>
A;Cross-references: UNIPROT:Q91317; GB:AE004598; NID:g9947687; PIDN:AA0510
A;Experimental source: strain PAO1
C;Genetics:
A;Gene: PscE; PA1718

Query Match 14.0%; Score 6; DB 2; Length 67;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 13 TALETR 18
|||||
Db 3 TALETR 8

RESULT 13
AE1003
hypothetical protein STY4338 [imported] - Salmonella enterica subsp. enterica serovar Typhimurium
C;Species: Salmonella enterica subsp. enterica serovar Typhimurium
A;Note: this species has also been called Salmonella typhimurium
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C;Accession: AE1003
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moller, S.; O'Gaora, P.
A;Reference number: AB0502; MUID:21534947; PMID:1167608
A;Accession: AE1003
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-74 <PAR>
A;Cross-references: GB:AL513382; PIDN:CAD08153.1; PID:gl6505129; GSPDB:GN00176
C;Genetics:
A;Gene: STY4338

Query Match 14.0%; Score 6; DB 2; Length 74;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 24 QRVQL 29
|||||
Db 12 QRVQL 17

RESULT 14
AE1953
SLYX protein homolog NMA0371 [imported] - Neisseria meningitidis (strain Z2491 serogroup B)
C;Species: Neisseria meningitidis
C;Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 09-Jul-2004
C;Accession: AE1953
R;Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel, J.; Holroyd, S.; Jørgensen, K.; Leather, S.; Mølle, S.; Mungall, K.; Quail, M.A.; Rajandream, A.; Rasmussen, M.; Saunders, N.J.; Seeger, A.; Skelton, J.; Stevens, K.; White, O.; Wren, B.; Young, P.; Berrington, J.; Brown, J.; Chillingworth, T.; Churcher, C.; Davis, A.; De Lencastre, E.; et al.
A;Title: Complete genome sequence of a serogroup A strain of Neisseria meningitidis Z2491.
A;Reference number: AB1775; MUID:20222556; PMID:10761919
A;Accession: AE1953
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-74 <PAR>
A;Cross-references: UNIPROT:Q9JWH4; GB:AL162753; GB:AL157959; NID:g7379120; PIDN:CAB8367
A;Experimental source: serogroup A, strain Z2491
C;Genetics:
A;Gene: slyX; NMA0371

Query Match 14.0%; Score 6; DB 2; Length 74;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 VAEQR 24
|||||
Db 32 VAEQR 37

RESULT 15
C81010
slyX protein, probable NMB2063 [imported] - Neisseria meningitidis (strain MCS8 serogroup B)
C;Species: Neisseria meningitidis
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C;Accession: C81010
R;Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.; Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.; et al.
A;Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MCS8.
A;Reference number: AB1000; MUID:20175755; PMID:10710307

A;Accession: C81010
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-74 <TET>
A;Cross-references: UNIPROT:Q9JXG3; GB:AE002556; GB:AE002098; NID:G7227322; PIDN:AAF4238
A;Experimental source: serogroup B, strain MC58
C;Genetics:
A;Gene: NMB2063

Query Match 14.0%; Score 6; DB 2; Length 74;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 VAE LRQ 24
| | | | |
Db 32 VAE LRQ 37

Search completed: July 26, 2005, 12:20:50
Job time : 20.7083 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 26, 2005, 11:59:44 ; Search time 97.3472 Seconds
(without alignments)
226.194 Million cell updates/sec

Title: US-10-624-218-3
Perfect score: 43
Sequence: 1 LEIEAFLERNTALETRVA.....QRVQLNRVSQVTRYGPL 43

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 1612378 seqs, 512079187 residues

Word size : 0
Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : UniProt_03.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	9	20.9	231	2 Q8LLY1	Q8LLY1 oryza sativ
2	9	20.9	601	2 Q8S853	Q8S853 oryza sativ
3	9	20.9	1106	2 Q72R15	Q72R15 leptospira
4	9	20.9	1106	2 Q8F4S3	Q8F4S3 leptospira
5	8	18.6	186	2 Q334S3	Q334S3 thiobacillu
6	8	18.6	293	2 Q57673	Q57673 brachydanio
7	8	18.6	300	2 Q9W722	Q9W722 anopheles g
8	8	18.6	395	2 Q7PYD2	Q7PYD2 anopheles g
9	8	18.6	449	2 Q6ATS2	Q6ATS2 oryza sativ
10	8	18.6	455	2 Q9X2P3	Q9X2P3 bacteroides
11	8	18.6	520	2 Q9VK31	Q9VK31 drosophila
12	8	18.6	524	2 Q6UUJ4	Q6UUJ4 oryza sativ
13	8	18.6	766	2 Q7X7W5	Q7X7W5 oryza sativ
14	8	18.6	839	2 Q84T33	Q84T33 oryza sativ
15	8	18.6	858	2 Q8LM25	Q8LM25 oryza sativ
16	8	18.6	903	2 Q6UUB0	Q6UUB0 oryza sativ
17	8	18.6	959	2 Q7XNM6	Q7XNM6 oryza sativ
18	8	18.6	1041	2 Q7UGM9	Q7UGM9 rhodopirell
19	8	18.6	1077	2 Q8LM33	Q8LM33 oryza sativ
20	8	18.6	1094	2 Q8S682	Q8S682 oryza sativ
21	8	18.6	1094	2 Q8S689	Q8S689 oryza sativ
22	8	18.6	1094	2 Q7G776	Q7G776 oryza sativ
23	8	18.6	1094	2 Q7G783	Q7G783 oryza sativ
24	8	18.6	1141	2 Q7XRZ0	Q7XRZ0 oryza sativ
25	8	18.6	1149	2 Q75HC5	Q75HC5 oryza sativ
26	8	18.6	1162	2 Q9I3I6	Q9I3I6 pseudomonas
27	8	18.6	1174	2 Q8S6M3	Q8S6M3 oryza sativ
28	8	18.6	1174	2 Q7XF67	Q7XF67 oryza sativ
29	8	18.6	1271	2 Q7XW55	Q7XW55 oryza sativ
30	8	18.6	1309	2 Q75H15	Q75H15 oryza sativ
31	8	18.6	1316	2 Q65WZ9	Q65WZ9 oryza sativ

32 8 18.6 1372 2 Q6ASW9 Q6ASW9 oryza sativ
33 8 18.6 1376 2 Q7XQ55 Q7XQ55 oryza sativ
34 8 18.6 1383 2 Q8L592 Q8L592 oryza sativ
35 8 18.6 1418 2 Q6AVD1 Q6AVD1 oryza sativ
36 8 18.6 1436 2 Q7FAL2 Q7FAL2 oryza sativ
37 8 18.6 1490 2 Q65WZ2 Q65WZ2 oryza sativ
38 8 18.6 1672 2 Q6UUS9 Q6UUS9 oryza sativ
39 8 18.6 1823 2 Q7XNM4 Q7XNM4 oryza sativ
40 7 16.3 180 2 Q6P707 Q6P707 homo sapien
41 7 16.3 184 2 Q6Y396 Q6Y396 taro bacill
42 7 16.3 185 2 Q7WYA9 Q7WYA9 pseudomonas
43 7 16.3 185 2 Q83XB3 Q83XB3 pseudomonas
44 7 16.3 186 1 TNR2_ECOLI P04130 escherichia
45 7 16.3 186 2 Q67JG1 Q67JG1 symbiobacte

ALIGNMENTS

RESULT 1
Q8LLY1 ID Q8LLY1 PRELIMINARY; PRT; 231 AA.
AC Q8LLY1; 01-OCT-2002 (Tremblrel. 22, Created)
DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
DT 05-JUL-2004 (Tremblrel. 27, Last annotation update)
DE Putative retroelement.
GN ORFNames=OSJNAA0053D03.7;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoidae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Wing R.A., Yu Y., Soderlund C., Kim H.-R., Rambo T., Sasaki C.,
RA Currie J., Collura K.;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA The Rice Chromosome 10 Sequencing Consortium;
RT "In-depth view of structure, activity, and evolution of rice
RT chromosome 10.";
RL Science 300:1566-1569(2003).
RN [3]
RP SEQUENCE FROM N.A.
RA Buell C.R., Wing R.A., McCombie W.R., Messing J., Yuan Q.;
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC131968; AAN04942.1; -;
DR EMBL; AB017074; AAP52923.1; -;
DR Gramene; Q8LLY1; -;
SQ SEQUENCE 231 AA; 26041 MW; A1E1C6D413209C73 CRC64;

Query Match 20.9%; Score 9; DB 2; Length 231;
Best Local Similarity 100.0%; Pred.No. 0.75;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 25 RVQRLNRV 33
| | | | | | | | | |
Db 99 RVQRLNRV 107
| | | | | | | | | |

RESULT 2
Q8S853 ID Q8S853 PRELIMINARY; PRT; 601 AA.
AC Q8S853; 01-JUN-2002 (Tremblrel. 21, Created)
DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE Hypothetical protein OSJNBA0053D03.3.
GN Name=OSJNBA0053D03.3;
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

```

OC  Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC  Ehrhartoideae; Oryzaceae; Oryza.
OX  NCBI_TaxID=4530;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=Nipponbare;
RA  McCombie W.R., Cordum H., Minx P., de la Bastide M., Spiegel L.,
RA  Nascimento L., Zutavern T., Ballja V., Bell M., Baker J., Santos L.,
RA  Miller B., Katzenberger F., Muller S., Preston R., Kirchoff K.,
RA  Kuit K., King L., Yang C., O'Saughnessy A., Palmer L., Dedhia N.,
RA  Wilson R.;
RL  Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR  EMBL; AC021892; AAM00939.1; -.
DR  Gramene; Q88583; -.
DR  InterPro; IPR008916; Retrov_capsid_C.
DR  Hypothetical protein.
KW  Hypothetical protein.
SQ  SEQUENCE 601 AA; 66130 MW; ACE70BC38C1C1590 CRC64;

Query Match          20.9%; Score 9; DB 2; Length 601;
Best Local Similarity 100.0%; Pred. No. 1.7; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0;

OY  25 RVQRLNRV 33
DB  469 RVQRLNRV 477

RESULT 3
ID  Q72R15      PRELIMINARY;      PRT; 1106 AA.
AC  Q72R15;
DT  05-JUL-2004 (TrEMBLrel. 27, Created)
DT  05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT  05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE  Heavy metal efflux pump.
GN  Name=czcA; OrderedLocusNames=LIC11937;
OS  Leptospira interrogans (serogroup Icterohaemorrhagiae / serovar
OS  Copenhagen).
OC  Bacteria; Spirochaetes; Spirochaetales; Leptospiaceae; Leptospira.
OC  NCBI_TaxID=44275;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=Fiocruz LI-130;
RX  PubMed=15028702; DOI=10.1128/JB.186.7.2164-2172.2004;
RA  Nascimento A.L.T.O., Ko A.I., Martins E.A.L., Monteiro-Vitorello C.B.,
RA  Ho P.L., Haake D.A., Verjovski-Almeida S., Hartskeerl R.A.,
RA  Marques M.V., Oliveira M.C., Menck C.F.W., Leite L.C.C., Carrer H.,
RA  Coutinho L.L., Degraeve W.M., Dellagostin O.A., El-Dorry H.,
RA  Ferro E.S., Ferro M.I.T., Furlan L.R., Gamberini M., Gigliotti E.A.,
RA  Goes-Neto A., Goldman G.H., Goldman M.H.S., Harakava R.,
RA  Jeronimo S.M.B., Junqueira-de-Azevedo J.L.M., Kimura E.T.,
RA  Kuramae E.B., Lemos E.G.M., Lemos M.V.F., Marino C.L., Nunes L.R.,
RA  de Oliveira R.C., Pereira G.G., Reis M.S., Schriefer A.,
RA  Siqueira W.J., Sommer P., Tsai S.M., Simpson A.J.G., Ferro J.A.,
RA  Camargo L.E.A., Kitajima J.P., Setubal J.C., Van Sluys M.A.;
RT  "Comparative genomics of two Leptospira interrogans serovars reveals
RT  novel insights into physiology and pathogenesis.";
RL  J. Bacteriol. 186:2164-2172(2004).
DR  EMBL; AE017294; AAS70519.1; -.
DR  GO; GO:0016021; C:integral to membrane; IEA.
DR  GO; GO:0008324; F:cation transporter activity; IEA.
DR  GO; GO:0006812; P:cation transport; IEA.
DR  InterPro; IPR001036; Acrflvin_res.
DR  InterPro; IPR004763; Czca.
DR  PRINTS; PR00702; ACRIFLAVINRP.
DR  TIGRFAMs; TIGR00914; 2A0601; 1.
KW  Complete proteome.
SQ  SEQUENCE 1106 AA; 123454 MW; E51315DD59F8BE47 CRC64;

Query Match          20.9%; Score 9; DB 2; Length 1106;
Best Local Similarity 100.0%; Pred. No. 2.9; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0;

OY  19 VAELRQRVQ 27
DB  853 VAELRQRVQ 861

RESULT 4
ID  Q8F4S3      PRELIMINARY;      PRT; 1106 AA.
AC  Q8F4S3;
DT  01-MAR-2003 (TrEMBLrel. 23, Created)
DT  01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT  01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE  Cation efflux system protein.
GN  Name=czcA3; OrderedLocusNames=LAI967;
OS  Leptospira interrogans.
OC  Bacteria; Spirochaetes; Spirochaetales; Leptospiaceae; Leptospira.
OC  NCBI_TaxID=173;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;
RX  MEDLINE=22598143; PubMed=12712204; DOI=10.1038/nature01597;
RA  Ren S.-X., Fu G., Jiang X.-G., Zeng R., Miao Y.-G., Xu H.,
RA  Zhang Y.-X., Xiong H., Lu G., Lu L.-F., Jiang H.-Q., Jia J., Tu Y.-F.,
RA  Jiang J.-X., Gu W.-Y., Zhang Y.-Q., Cai Z., Sheng H.-H., Yin H.-F.,
RA  Zhang Y., Zhu G.-P., Wan M., Huang H.-L., Qian Z., Wang S.-Y., Ma W.,
RA  Yao Z.-J., Shen Y., Qiang B.-Q., Xia Q.-C., Guo X.-K., Danchin A.,
RA  Saint Girons I., Somerville R.L., Wen Y.-M., Shi M.-H., Chen Z.,
RA  Xu J.-G., Zhao G.-P.;
RT  "Unique physiological and pathogenic features of Leptospira
RT  interrogans revealed by whole-genome sequencing.";
RL  Nature 422:888-893(2003).
DR  EMBL; AE011369; AAN49166.1; -.
DR  GO; GO:0016021; C:integral to membrane; IEA.
DR  GO; GO:0008324; F:cation transporter activity; IEA.
DR  GO; GO:0006812; P:cation transport; IEA.
DR  InterPro; IPR001036; Acrflvin_res.
DR  InterPro; IPR004763; Czca.
DR  Pfam; PF00873; ACR tran; 1.
DR  PRINTS; PR00702; ACRIFLAVINRP.
DR  TIGRFAMs; TIGR00914; 2A0601; 1.
KW  Complete proteome.
SQ  SEQUENCE 1106 AA; 123374 MW; 4FB761106F693BEF CRC64;

Query Match          20.9%; Score 9; DB 2; Length 1106;
Best Local Similarity 100.0%; Pred. No. 2.9; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0;

OY  19 VAELRQRVQ 27
DB  853 VAELRQRVQ 861

RESULT 5
ID  Q934S3      PRELIMINARY;      PRT; 186 AA.
AC  Q934S3;
DT  01-DEC-2001 (TrEMBLrel. 19, Created)
DT  01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT  01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE  Resolvase.
GN  Name=cnpR;
OS  Thiobacillus ferrooxidans.
OC  Bacteria; Proteobacteria; Gammaproteobacteria; Acidithiobacillales;
OC  Acidithiobacillaceae; Acidithiobacillus.
OX  NCBI_TaxID=920;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=G66;
RA  Kalyaeva E.S., Kholodii G.Y., Bass I.A., Gorlenko A.M., Yurieva O.V.,
RA  Nikiforov V.G.;
RT  "Tn5037, a Tn21-like mercury resistance transposon from Thiobacillus
RT  ferrooxidans.";
RL  Russ. J. Genet. 37:972-975(2001).

```

```
DR EMBL; AJ251743; CAC69253.1; -.
DR HSPSP; P03012; 2RSL.
DR GO; GO:0000150; F:recombinase activity; IEA.
DR GO; GO:0006310; P:DNA recombination; IEA.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR009057; Homeodomain_like.
DR InterPro; IPR006120; HTH_7.
DR InterPro; IPR006118; Recombinase_S.
DR InterPro; IPR006119; resolvase_N.
DR Pfam; PF02796; HTH_7; 1.
DR Pfam; PF00239; Resolvase; 1.
DR PROSITE; PS00018; EF_HAND; UNKNOWN 1.
DR PROSITE; PS00397; RECOMBINASES_1; 1.
DR PROSITE; PS00398; RECOMBINASES_2; 1.
DR PROSITE; PS00399; RECOMBINASES_3; 1.
SQ SEQUENCE 186 AA; 21291 MW; 265D8E220EA38281 CRC64;

Query Match 18.6%; Score 8; DB 2; Length 186;
Best Local Similarity 100.0%; Pred. No. 6.3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 RVAELRQR 25
Db 150 RVAELRQR 157

RESULT 6
O57673 PRELIMINARY; PRT; 293 AA.
AC O57673;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Thytrotroph embryonic factor beta.
GN Name=tef; Synonyms=tefbeta;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98094234; PubMed=9434145; DOI=10.1016/S0167-4781(97)00150-4;
RU Xu X., Liu L., Wong K.C., Ge R.;
RT "Cloning and characterization of two isoforms of the zebrafish
thyrotroph embryonic factor.";
RL Biochim. Biophys. Acta 1395:13-20(1998).
CC -!- SIMILARITY: Belongs to the bZIP family.
DR EMBL; U43671; AAC00206.1; -.
DR ZFIN; ZDB-GENE-990415-264; tef.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR004827; TF_bZIP.
DR SMART; SM00338; BRLZ; 1.
DR PROSITE; PS50217; BZIP; 1.
DR DNA-binding; Nuclear protein.
KW DNA-binding; Nuclear protein.
SQ SEQUENCE 300 AA; 33522 MW; 2680CF628ED15EB1 CRC64;

Query Match 18.6%; Score 8; DB 2; Length 300;
Best Local Similarity 100.0%; Pred. No. 9.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 AAFLEREN 12
Db 262 AAFLEREN 269

RESULT 8
Q7PYD2 PRELIMINARY; PRT; 395 AA.
AC Q7PYD2;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE AGCP12016.
GN Name=agCGS3582; ORFNames=ENSANGG00000018310;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAINS=PEST;
RA Anopheles Genome Sequencing Consortium;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the bZIP family.
CC -!- CAUTION: The sequence shown here is derived from an
EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
preliminary data.
DR EMBL; AAAB01008987; EAA01664.1; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR004827; TF_bZIP.
DR PROSITE; PS50217; BZIP; 1.
DR PROSITE; PS00036; BZIP_BASIC; 1.
DR DNA-binding; Nuclear protein.
KW DNA-binding; Nuclear protein.
SQ SEQUENCE 395 AA; 43154 MW; 8A9870E4982A274F CRC64;

Query Match 18.6%; Score 8; DB 2; Length 395;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 AAFLEREN 12
```

```
DR EMBL; AJ251743; CAC69253.1; -.
DR HSPSP; P03012; 2RSL.
DR GO; GO:0000150; F:recombinase activity; IEA.
DR GO; GO:0006310; P:DNA recombination; IEA.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR009057; Homeodomain_like.
DR InterPro; IPR006120; HTH_7.
DR InterPro; IPR006118; Recombinase_S.
DR InterPro; IPR006119; resolvase_N.
DR Pfam; PF02796; HTH_7; 1.
DR Pfam; PF00239; Resolvase; 1.
DR PROSITE; PS00018; EF_HAND; UNKNOWN 1.
DR PROSITE; PS00397; RECOMBINASES_1; 1.
DR PROSITE; PS00398; RECOMBINASES_2; 1.
DR PROSITE; PS00399; RECOMBINASES_3; 1.
SQ SEQUENCE 186 AA; 21291 MW; 265D8E220EA38281 CRC64;

Query Match 18.6%; Score 8; DB 2; Length 186;
Best Local Similarity 100.0%; Pred. No. 6.3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 RVAELRQR 25
Db 150 RVAELRQR 157

RESULT 6
O57673 PRELIMINARY; PRT; 293 AA.
AC O57673;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Thytrotroph embryonic factor beta.
GN Name=tef; Synonyms=tefbeta;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98094234; PubMed=9434145; DOI=10.1016/S0167-4781(97)00150-4;
RU Xu X., Liu L., Wong K.C., Ge R.;
RT "Cloning and characterization of two isoforms of the zebrafish
thyrotroph embryonic factor.";
RL Biochim. Biophys. Acta 1395:13-20(1998).
CC -!- SIMILARITY: Belongs to the bZIP family.
DR EMBL; U43671; AAC00206.1; -.
DR ZFIN; ZDB-GENE-990415-264; tef.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR004827; TF_bZIP.
DR SMART; SM00338; BRLZ; 1.
DR PROSITE; PS50217; BZIP; 1.
DR PROSITE; PS00036; BZIP_BASIC; 1.
DR DNA-binding; Nuclear protein.
KW DNA-binding; Nuclear protein.
SQ SEQUENCE 300 AA; 33522 MW; 2680CF628ED15EB1 CRC64;

Query Match 18.6%; Score 8; DB 2; Length 300;
Best Local Similarity 100.0%; Pred. No. 9.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 AAFLEREN 12
Db 249 AAFLEREN 256

RESULT 7
Q9W722 PRELIMINARY; PRT; 300 AA.
AC Q9W722;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
```

DR InterPro: IPR001668; Mob Pre.
DR Pfam: PF01076; Mob Pre: 1.
SQ SEQUENCE 455 AA; 52650 MW; 35AE86CC18C02AAA CRC64;

Query Match 18.6%; Score 8; DB 2; Length 455;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 LERENTAL 15
DB 278 LERENTAL 285

RESULT 11
Q9VK31 PRELIMINARY; PRT; 520 AA.
AC Q9VK31;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 25-OCT-2004 (TReMBLrel. 28, Last annotation update)
DE CG5439-PA (ID35592p).
GN ORFNames=CG5439;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Anantides P.G., Scher S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brattton R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Gabor G.L.,
RA Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bertram B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Folsler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwan C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacht J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.Y., Wassarman D.A., Weissstock G.M., Weissbach J.,
RA Williams S.M., Woodage, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhou X., Zhou L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426065; PubMed=12537568;
RA Celnik S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,

DB 362 AAFLEREN 369

Q6ATS2 PRELIMINARY; PRT; 449 AA.
AC Q6ATS2;
DT 25-OCT-2004 (TReMBLrel. 28, Created)
DT 25-OCT-2004 (TReMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TReMBLrel. 28, Last annotation update)
DE Hypothetical protein OSJNB0056010.19 (Fragment).
GN Name=OSJNB0056010.19;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Buell C.R., Yuan Q., Ouyang S., Liu J., Gansberger K., Jones K.M.,
RA Overton II L.L., Taitrin T., Kim M.M., Bera J.J., Jin S.S.,
RA Fadrosch D.W., Tallon L.J., Koo H., Zismann V., Hsiao J., Blunt S.,
RA Vanaken S.S., Riedmuller S.B., Utterback T.T., Feldblyum T.V.,
RA Yang Q.Q., Haas B.J., Suh B.B., Peterson J.J., Quackenbush J.,
RA White O., Salzberg S.L., Fraser C.M.;
RT "Oryza sativa chromosome 3 BAC OSJNB0056010 genomic sequence.";
RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Buell R.;
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC135564; AAT81757.1; -;
DR InterPro: IPR009007; Pept_Aspartic.
KW Hypothetical protein.
FT NON TER 449
SQ SEQUENCE 449 AA; 50837 MW; 598354E253C6C1A CRC64;

Query Match 18.6%; Score 8; DB 2; Length 449;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 RVQRLNR 32
DB 165 RVQRLNR 172

RESULT 10
Q9X2P3 PRELIMINARY; PRT; 455 AA.
AC Q9X2P3;
DT 01-NOV-1999 (TReMBLrel. 12, Created)
DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Mobilization protein BmpH.
GN Name=bmpH;
OS Bacteroides fragilis.
OC Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;
OC Bacteroidaceae; Bacteroides.
OX NCBI_TaxID=817;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LV23; TRANSPOSON=Tn5520;
RX MEDLINE=98214109; PubMed=10198023;
RA Vedantam G., Novicki T.O., Hecht D.W.;
RT "Bacteroides fragilis transfer factor Tn5520: the smallest bacterial
RT mobilizable transposon containing single integrase and mobilization
RT genes that function in Escherichia coli.";
RL J. Bacteriol. 181:2564-2571(1999).
DR EMBL; AF039866; AAD24590.1; -;
DR GO; GO:0005727; C:extrachromosomal circular DNA; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0006310; F:DNA recombination; IEA.

George R.A., Hoskins R.A., Lavery T., Muzny D.M., Nelson C.R., Pacle J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J., Swirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C., Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.; "Finishing a whole-genome shotgun: Release 3 of the Drosophila melanogaster euchromatic genome sequence.", *Genome Biol.* 3:RESEARCH0079-RESEARCH0079(2002).

[3]
SEQUENCE FROM N.A.
MEDLINE=22426070; PubMed=12537573;
Kaminker J.S., Bergman C.M., Krommiller B., Carlson J., Swirskas R., Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M., Ashburner M., Celniker S.E.; "The transposable elements of the Drosophila melanogaster euchromatin: a genome perspective.", *Genome Biol.* 3:RESEARCH0084-RESEARCH0084(2002).

[4]
SEQUENCE FROM N.A.
MEDLINE=22426069; PubMed=12537572;
Miara S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S., Hradscky P., Huang Y., Kaminker J.S., Millburn G.H., Prochownik S.E., Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P., Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A., Harris N.D., Richter J., Russo S., Schroeder A.J., Shu S.Q., Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M., Lewis S.E.; "Annotation of the Drosophila melanogaster euchromatic genome: a systematic review.", *Genome Biol.* 3:RESEARCH0083-RESEARCH0083(2002).

[5]
SEQUENCE FROM N.A.
FlyBase;
Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.

[6]
SEQUENCE FROM N.A.
FlyBase;
Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.

[7]
SEQUENCE FROM N.A.
STRAIN=Berkeley;
Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J., Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E., George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G., Miranda A., Mungall C.J., Nunoo J., Pacle J., Paragas V., Park S., Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M., Celniker S.; Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AE003639; AAF53251.1; -;
DR EMBL; AY095047; AAM11375.1; -;
DR GO; 0007242; P:intracellular signaling cascade; IEA.
DR InterPro; IPR001683; PX.
DR InterPro; IPR004012; RUN.
DR InterPro; IPR001680; WD40.
DR Pfam; PF00787; PX; 1.
DR Pfam; PF02759; RUN; 1.
DR SMART; SM00312; PX; 1.
DR SMART; SM00593; RUN; 1.
DR PROSITE; PS00195; PX; 1.
DR PROSITE; PS50826; RUN; 1.
DR PROSITE; PS00678; WD REPEATS 1; UNKNOWN 1.
SQ SEQUENCE 520 AA; 59256 MW; 2BA4F6EBB39AC605 CRC64;

Query Match 18.6%; Score 8; DB 2; Length 520;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 LSTRVAEL 22
|||||

Db 367 LSTRVAEL 374

RESULT 12

Q6UUJ4 PRELIMINARY; PRT; 524 AA.
AC Q6UUJ4
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN ORFNames=OSJNBa0017M13.28, OSJNBa0095C12.2;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=14716315; DOI=10.1038/ng1289;
RA Nagaki K., Cheng Z., Ouyang S., Talbert P.B., Kim M., Jones K.M., Henikoff S., Buell C.R., Jiang J.;
RT "Sequencing of a rice centromere uncovers active genes."; *Nat. Genet.* 36:138-145(2004).
RL EMBL; AY360386; AAQ56358.1; -;
DR EMBL; AY360385; AAQ56325.1; -;
DR InterPro; IPR005162; Retrotrans gag.
DR InterPro; IPR008916; Retrov capsid_C.
DR Pfam; PF03732; Retrotrans_gag; 1.
KW Hypothetical protein.
SQ SEQUENCE 524 AA; 59956 MW; FCB711C42C2AEC89 CRC64;

Query Match 18.6%; Score 8; DB 2; Length 524;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 25 RVQRLNR 32
|||||

Db 322 RVQRLNR 329

RESULT 13
Q7X7W5 PRELIMINARY; PRT; 766 AA.
AC Q7X7W5
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE OSJNB0046P18.13 protein (OSJNBa0017P10.8 protein).
GN Name=OSJNB0046P18.13; Synonyms=OSJNBa0017P10.8;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=12447439; DOI=10.1038/nature01183;
RA Feng Q., Zhang Y., Hao P., Wang S., Fu G., Huang Y., Li Y., Zhu J., Liu Y., Hu X., Jia P., Zhang Y., Zhao Q., Ying K., Yu S., Tang Y., Wang Q., Zhang L., Lu Y., Mu J., Zhang L.S., Yu Z., Fan D., Liu X., Lu T., Li C., Wu Y., Sun T., Lei H., Li T., Hu H., Guan J., Wu M., Zhang R., Zhou B., Chen Z., Jin Z., Wang R., Yin H., Cai Z., Ren S., Lv G., Gu W., Zhu G., Tu Y., Jia J., Zhang Y., Chen J., Kang H., Chen X., Shao C., Sun Y., Hu Q., Zhang X., Zhang W., Wang L., Ding C., Sheng H., Gu J., Chen S., Ni L., Zhu F., Chen W., Lan L., Lai Y., Cheng Z., Gu M., Jiang J., Li J., Hong G., Xue Y., Han B.;
RT "Sequence and analysis of rice chromosome 4."; *Nature* 420:316-320(2002).
RL EMBL; AL606630; CAE03297.2; -;
DR EMBL; AL731578; CAE04931.2; -;
DR Gramene; Q7X7W5; -;
DR InterPro; IPR011036; PH related.
DR InterPro; IPR000477; RVTse.
DR Pfam; PF00078; RVT_1; 1.
KW RNA-directed DNA polymerase; Transferase.

```
SQ SEQUENCE 766 AA; 87698 MW; E4B3A4F8D04882F8 CRC64;
Query Match 18.6%; Score 8; DB 2; Length 766;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 RVORLRNR 32
Db 252 RVORLRNR 259

RESULT 14
Q84T33 PRELIMINARY; PRT; 839 AA.
AC Q84T33;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Putative GAG-POL (Putative reverse transcriptase).
GN Names=OSJNB0031104.34; Synonyms=OJ1324.A07.7;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OC NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Buell C.R., Yuan Q., Ouyang S., Liu J., Gansberger K., Jones K.M.,
RA Overton II L.L., Taitrin T., Kim M.M., Bera J.J., Jin S.S.,
RA Fadrosch D.W., Tallon L.J., Koo H., Zismann V., Heiao J., Blunt S.,
RA Vanaken S.S., Riedmuller S.B., Utterback T.T., Feldblyum T.V.,
RA Yang Q.Q., Haas B.J., Suh B.B., Peterson J.J., Quackenbush J.,
RA White O., Salzberg S.L., Fraser C.M.;
RA Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Buell R.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Buell C.R., Yuan Q., Ouyang S., Liu J., Gansberger K., Jones K.M.,
RA Overton II L.L., Taitrin T., Kim M.M., Bera J.J., Jin S.S.,
RA Fadrosch D.W., Tallon L.J., Koo H., Zismann V., Heiao J., Blunt S.,
RA Vanaken S.S., Riedmuller S.B., Utterback T.T., Feldblyum T.V.,
RA Yang Q.Q., Haas B.J., Suh B.B., Peterson J.J., Quackenbush J.,
RA White O., Salzberg S.L., Fraser C.M.;
RT "Oryza sativa chromosome 3 BAC OJ1324.A07 genomic sequence.";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Buell R.;
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC135560; AA066577.1; -.
DR EMBL; AC115688; AAT77827.1; -.
DR Gramene; Q84T33; -.
DR GO; GO:0004523; F:ribonuclease H activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.
DR InterPro; IPR002156; RNaseH.
DR InterPro; IPR000477; RVTse.
DR Pfam; PF00075; RnaseH; 1.
DR Pfam; PF00078; RVT_1; 1.
KW RNA-directed DNA polymerase; Transferase.
SQ SEQUENCE 839 AA; 95600 MW; AFEF90C0188E88A6 CRC64;

Query Match 18.6%; Score 8; DB 2; Length 839;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 RVORLRNR 32
Db 252 RVORLRNR 259

us-10-624-218-3-olig.rup

SQ SEQUENCE 766 AA; 87698 MW; E4B3A4F8D04882F8 CRC64;
Query Match 18.6%; Score 8; DB 2; Length 766;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 RVORLRNR 32
Db 252 RVORLRNR 259

RESULT 15
Q8LM25 PRELIMINARY; PRT; 858 AA.
AC Q8LM25;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Putative retroelement.
GN ORFNAMES=OSUNAA0019N10.17;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OC NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA McCombie W.R., de la Bastide M., Spiegel L., Preston R., Ferraro K.,
RA Kuit K., Nascimento L., Zutavern T., Balija V., Bell M., Baker J.,
RA Miller B., Katzenberger F., Muller S., King L., Sullivan P., Yang C.,
RA Dike S., O'Shaughnessy A., Palmer L., Dedhia N.;
RA Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA The Rice Chromosome 10 Sequencing Consortium;
RA "In-depth view of structure, activity, and evolution of rice
RT chromosome 10.";
RL Science 300:1566-1569 (2003).
RN [3]
RP SEQUENCE FROM N.A.
RA Buell C.R., Wing R.A., McCombie W.R., Messing J., Yuan Q.;
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC124213; AAM74459.1; -.
DR EMBL; AE017087; AAP53484.1; -.
DR Gramene; Q8LM25; -.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.
DR InterPro; IPR009007; Pept_Aspartic.
DR InterPro; IPR011036; PH_related.
DR Pfam; PF00078; RVT_1; 1.
KW RNA-directed DNA polymerase; Transferase.
SQ SEQUENCE 858 AA; 97208 MW; CC3277BA7BDA4410 CRC64;

Query Match 18.6%; Score 8; DB 2; Length 858;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 RVORLRNR 32
Db 297 RVORLRNR 304

Search completed: July 26, 2005, 12:19:39
Job time : 99.3472 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 27, 2005, 10:08:29 ; Search time 1780.92 Seconds
(without alignments)
1169.945 Million cell updates/sec

Title: US-10-624-218-3

Perfect score: 43

Sequence: 1 LEIEAAFLRENTALETRVA.....QRVQRLNRVSQVTRYGPL 43

Scoring table:

OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Word size: 1

Total number of hits satisfying chosen parameters: 9407124

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODE=frame+P2n.model -DEV=xlp
-Q/cgn2_1/USPTO_gpool_p/US10624218/runat_26072005_121435_3685/app_query.fasta_1.796
-DB=GenEmbl -QWMT=fastcap -SUFFIX=olip2n.rge -MINMATCH=0.1 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10624218@cgn 1.1 6148 @runat_26072005_121435_3685 -NCPU=6 -ICPU=3
-NO_WMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*

- 1: gb.ba.*
- 2: gb.hcg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.to.*
- 11: gb.sts.*
- 12: gb.sw.*
- 13: gb.un.*
- 14: gb.vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	9	20.9	11251	1	AE011369
2	9	20.9	150496	8	AC131968
3	9	20.9	156874	8	AP006237
4	9	20.9	175461	8	AC137747

5	9	20.9	199113	8	AC021892
6	9	20.9	232255	2	AC132702
7	9	20.9	300207	1	AE017294
8	9	20.9	304110	8	AE017074
9	9	20.9	305902	2	AC105818
10	8	18.6	440	11	G73580
11	8	18.6	520	8	WH7ALISA
12	8	18.6	891	6	AR505926
13	8	18.6	1459	5	DR096848
14	8	18.6	1594	3	AY060258
15	8	18.6	1620	6	AX755546
16	8	18.6	1659	6	C0590226
17	8	18.6	1659	8	AK071667
18	8	18.6	1682	3	ENRNAS
19	8	18.6	1751	9	BC015648
20	8	18.6	1796	6	C0583530
21	8	18.6	1825	3	AY095047
22	8	18.6	1867	8	AK106925
23	8	18.6	1950	8	AK109679
24	8	18.6	2000	6	AX756068
25	8	18.6	2716	9	BC025243
26	8	18.6	2750	5	DRU43671
27	8	18.6	2775	6	AX086074
28	8	18.6	3796	6	C0583529
29	8	18.6	3858	6	C0590225
30	8	18.6	3954	6	CQ737497
31	8	18.6	4576	8	AF291822
32	8	18.6	4692	1	AF038866
33	8	18.6	5186	6	C0590012
34	8	18.6	6155	6	C0583232
35	8	18.6	7771	1	TRE251743
36	8	18.6	13203	1	AB004581
37	8	18.6	24965	2	AC014507
38	8	18.6	32059	2	AC019906
39	8	18.6	36086	8	AC146915
40	8	18.6	38734	8	SPCC338
41	8	18.6	39200	9	AC024584
42	8	18.6	42380	8	AF005922
43	8	18.6	50913	7	AY129332
44	8	18.6	65115	8	AC135428
45	8	18.6	68992	8	AC120306

ALIGNMENTS

AE011369 11251 bp DNA linear BCT 28-APR-2003
Leptospiira interrogans serovar lai str. 56601 chromosome I, section 178 of 397 of the complete sequence.

AE011369 AE010300
AE011369.1 GI:24195644

Leptospiira interrogans serovar Lai str. 56601
Leptospiira interrogans serovar Lai str. 56601

Bacteria; Spirochaetes; Spirochaetales; Leptospiiraceae; Leptospiira.
1 (bases 1 to 11251)

Ren, S.X., Fu, G., Jiang, X.G., Zeng, R., Miao, Y.G., Xu, H., Zhang, Y.X.,
Xiong, H., Lu, L.F., Jiang, H.Q., Jia, J., Tu, Y.F., Jiang, J.X.,
Gu, W.Y., Zhang, Y.Q., Cai, Z., Sheng, H.H., Yin, H.F., Zhang, Y.,
Zhu, G.F., Wan, M., Huang, H.L., Qian, Z., Wang, S.Y., Ma, W., Yao, Z.J.,
Shen, Y., Qiang, B.Q., Xia, Q.C., Guo, X.K., Danchin, A., Saint
Girons, I., Somerville, R.L., Wen, Y.M., Shi, M.H., Chen, Z., Xu, J.G.
and Zhao, G.P.

Unique physiological and pathogenic features of Leptospiira
interrogans revealed by whole-genome sequencing

Nature 422 (6934), 888-893 (2003)
12712204

2 (bases 1 to 11251)

Ren, S., Fu, G., Jiang, X., Zeng, R., Xiong, H., Lu, L., Lu, G., Jiang, H.,
Ding, Y., Jia, J., Tu, Y., Gu, W., Cai, Z., Sheng, H., Yin, H., Zhang, Y.,
Zhu, G., Wang, S., Shen, Y., Qiang, B., Chen, Z., Wen, Y., Xu, J. and
Zhao, G.

TITLE Direct Submission
JOURNAL Submitted (12-MAR-2002) Chinese National Human Genome Center at Shanghai, 250 Bi Bo Road, Shanghai 201203, China
COMMENT Updated information will be available at our World Wide Web site (http://www.chgc.sh.cn/lep/). Comments to the authors are appreciated.

FEATURES
source 1. .11251
/organism="Leptospira interrogans serovar Lai str. 56601"
/mol_type="genomic DNA"
/strain="56601"
/serovar="Lai"
/db_xref="taxon:189518"
/chromosome="I"
/note="Strain 56601 is maintained by the Institute of Epidemiology and Microbiology, Chinese Academy of Preventive Medicine"
81. .458
/locus_tag="LA1962"
81. .458
/locus_tag="LA1962"
/codon_start=1
/transl_table=1
/product="hypothetical protein"
/protein_id="AA049161.1"
/db_xref="GI:24195645"
/translation="MKRWIHRFLILVLTAVILVPCEDGDTDLAFDVRGHEVHTLAKIRSQSHKTSQIETQSGIETDRLTIYTRGYDFDLDFVSGSFELSKVEVHYFSGKFNAPKTVFSLSSLLNLPPTV"
558. .2012
/locus_tag="LA1963"
558. .2012
/locus_tag="LA1963"
/codon_start=1
/transl_table=1
/product="outer membrane protein precursor CzcC"
/protein_id="AA049162.1"
/db_xref="GI:24195646"
/translation="MECTLNQQTILSGKPKKIFILLFFLVLITCVFSESKTEVLDSAIIPKNESSKSLNWEKPTFEKRVLDLKTAEKLRWNNLLLASRPNIDARVAGIEQGLVANNFIIDQSIYAEPTQYDFRSRGTVQIQOQLFLGKIGKGRVRAELNARMEQEFYDLARGVLTKLRFYAIYYRQAIISFYDQSLGALGKTYSAAEMGYKRRAILQAEVRLRLKFLFKGERDINILREADRLVLLNDLKNPKETKPIVEIPVEEALDLPEAKLKRELLVARNKPLDKKVAQALRYEANLELOHANAIPDLAAGPMYNRGTAQNYWCIITQALNVPIDFNQGNIKASEKAVLSKQELKNTLLENEVAVSIE TARVKEIDYKFKFTITREYIDLSKMLISYKRYITILEFADFFETYSRISIVEMLKLQTRMDAIEGVNFVSGVGGVLIPLKLKSAESGTENK"
2585. .3031
/locus_tag="LA1964"
2585. .3031
/locus_tag="LA1964"
/codon_start=1
/transl_table=1
/product="cation efflux system membrane protein"
/protein_id="AA049163.1"
/db_xref="GI:24195647"
/translation="MWFLNLFNLKKEKGRKVKVRFSSPNEWSGTALGDNDVP MTRVKKRIVISNAGYKLPKGMFANVPEPTANDTVVPPNAITVTEGGQYVFEVA PHEFVRREVTIGISTTRKVNVIETGTRKDRVVUGSILLKLSGFG"
3021. .3767
/locus_tag="LA1965"
3021. .3767
/locus_tag="LA1965"
/codon_start=1
/transl_table=1
/product="putative outer membrane protein"
/protein_id="AA049164.1"
/db_xref="GI:24195648"
/translation="WVFNNQKLVLFILCFNLVSVWAQSQKQSPSKDPPSPWEVYDPDKTKIKGEYTEDVLKSLPKSTNLFVVGASIGSPGKINILNLGYKDYKDVFRVSGKWNTHWGGQVDIGYGFPMKTPVTAHSVIVFGAFVDFPQFPIRGQGNQSPSGSS PFGYQHRDFTEDLIRSVITEQSPDLNLYLESRDRQKVLHNNQYVGLTYDFLLGN FFLQJGGGIGKGDYKPNQLLIQVGYLFDTR"

gene
CDS

gene
CDS

gene
CDS

gene
CDS

3773. .7102
/gene="czca2"
/locus_tag="LA1966"
3773. .7102
/gene="czca2"
/locus_tag="LA1966"
/codon_start=1
/transl_table=1
/product="cation efflux system protein"
/protein_id="AA049165.1"
/db_xref="GI:24195649"
/translation="MKFLNLLIESALRYRFTLLVAALLLAVGWSMLDIRKEAYSDIADTOVRLVAKPFGKATLEVERVTPMPTIERVLHSTPNLIVRRSRTINGLVVQFVFEEGTDDYFARMRLMEKVRDAVPEEVPTFLAPMSSPVGEYRVYVETAGTHTPMELRITIQDWIIVPKLQVSGIADVVTFGLPQKQHVTSPEKLIRYSYTVADVIEAISKNNLTGNGNFILOGSOLPIRSIGAIRTPEIDIEIVKTVANGVPVVRDIDSVIEISHPIPSGLGTVYVNEQGLIDVDSAVQGLVAMRWVEPNAGFGRVRAKVEIKDEINRYMPECTIRNTYDRGDLVYKTLRTVGTTLLEGIAVSVLSLVVIFIGLSKASIVVAVTIPFALLPFTMMNSSGISASLSLGLGADVIGIVDSAVVNMIRRYKNAIPAQKQGLVRTYOCATEVGT EVLFAILLIILAYLPIFSEFIEGRLEKPKMAFTLSFAIFGALLFTMTVVPVLMSPFYR RYFESKNPQPIEMHNVPYEWYERKYEKIVHYLVDRSKVMVIAFTGTLIVGGMSL GTEFLPSLDEGGFTLRLYPVPGISLPEAKFVKPIRGMIYKNEQVMILISQYGRNDG GTDPLPNRLEIYVGLKDYKNNEKITKEQLIRMNDELEGLPGVKVSFSQPIMDNLS GAIMGTIADLAVFVSGQDLKEMRHISQQLIDIVSKMGASVGLQECPAQLVRLK RAVAAARYGINISDIQNVLEAAVGMPEPISLYEGMDTTPKRALFGIAVRAKQYRQS ARETANIPIISPKGERIPLSELAETIQEDSPTMIPRODGKRTITVRLNVRGRDQGFV AELQKVKKEVKIPGFEVRYGQYENLARVKGQIAIVPTTIGIIFGLLPMYLNRLK SVIVALSCEPISLLGIVALLGRGYFNVSAVGVSIFSLGIATMAGILFVSKANLHR DDIMVSREASITLSAVSOLRPLKMLTLLALLGLIPATWASGVGSDVQRPPLATVMVGG L ASALFLVLTVMPSLYILVMGSKEDVYKSGSSKSLSGDYASIVPEYFISDDEYQA PTEKSKKLLKNKKKKKKI"
7637. .10957
/gene="czca3"
/locus_tag="LA1967"
7637. .10957
/gene="czca3"
/locus_tag="LA1967"
/codon_start=1
/transl_table=1
/product="cation efflux system protein"
/protein_id="AA049166.1"
/db_xref="GI:24195650"
/translation="MIRNLGEMRFRVATLAASVAVFGIWMADIRKEAYSIDIATQVRLIAPFGKAAVEEVERVTPIERVLNAPKVAVRSRTINGLVVQFVFEEDGTDYFATRLRLERVDADIPSEVETPLGPMSSPVEIYRVYVESKANHMTMELRTIQDWN IMPKMSIPIGIADVTFGLPQKQHVTTDLRIYKLTIEDVINAIOKNNLTGNNL LIQOEGGPIRSIGAIRTPEHTENIVKTVANGVPVVRDIDSVIEISHPIPSGLVGTI QNDQGLIDVDSVQGLVARIWGDNPINMERIRAKVKEINDNLYPDGVQIRNTYDRT DLVNTLRTIGKTLVEGVVVSVLIFFTIGSIKASVVVATIPFAMLPFLMNTIGI PASLSLGLADFGIIVDGAIVNMVNRMYRDSPTDKKGIIVRTVDAASEVGTETI FSLIILAYLPFISFERIEGRLEKPKMAFTISFAIFGALLIFSM TALPVLMSIYRNYF EQNPQPIEWHNPFYDMLERKYEKIVHYLVDRSKVMVIAFTGTLIVGGMSLGT EFLPMDEMGFNLRIPFPVGSILPESKGFIPKIREMIYNEQVNVLSQIRGNDGDDTP LPPNRLVLSLKYDNWKEKITKQELIRMNDELEAGLPGARVSFSQPIMDNLS EAI MGTIADLAVFVSGQDLKEMRQITASEVLEIVKMGASVGLQECPAQLVRLK RAAAGIYVNDVQVMVEAALCMQRIDTLYEGSDVPKPARFGIVVRFSKVRSSKRA IENIPITISPKGERIPLSELAETIQEDSPTMIPRODGKRTITVRLNVRGRDQGFV RORVQKKILPGEYEIRYGGYENLARVKGQIAIVPTTIGIIFGLLPMYLNRLK KYV VALACIPSLVGVIVALLRGYCFNVSGGVGFSIFSLGIATMAGVLFVSRNHLHEDD ETSVREAAKAAVIOQLRPMLTILLALLGLIPATLASGVGSDVQRPPLATVIVGLFSA LLALVTLPSLYILVNGERDYPVKQNLPEYSLDQYPLEEVESELDVSTETKSSN LGKEKTSKDKTKKKKKK"

ORIGIN

Alignment Scores:
Pred. No.: 21.6 Length: 11251
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 20.93% Indels: 0
DB: 1 Gaps: 0
US-10-624-218-3 (1-43) x AE011369 (1-11251)

```
Oy 19 ValAlaGluLeuArgGlnArgValGln 27
|||||
Db 10193 GTACGGAGCTAAGACAAAGATTTCAG 10219

RESULT 2
AC131968/c
LOCUS
DEFINITION
Oryza sativa (japonica cultivar-group) chromosome 10 clone
OSJNAa0053D03, complete sequence.
AC131968
AC131968.1 GI:22532547
HTG.
Oryza sativa (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzoideae; Oryza.
1 (bases 1 to 150496)
Wing, R.A., Yu, Y., Soderlund, C., Kim, H.-R., Rambo, T., Saeki, C.,
Currie, J. and Collura, K.
Rice Genomic Sequence
Unpublished
2 (bases 1 to 150496)
Wing, R.A., Yu, Y., Soderlund, C., Kim, H.-R., Rambo, T., Saeki, C.,
Currie, J. and Collura, K.
Direct Submission
Submitted (28-AUG-2002) Arizona Genomics Institute, University of
Arizona, 303 Forbes, Tucson, AZ 85721, USA
3 (bases 1 to 150496)
Wing, R.A., Yu, Y., Soderlund, C., Kim, H.-R., Rambo, T., Saeki, C.,
Currie, J., Collura, K. and Thurmond, S.K.
Direct Submission
Submitted (05-SEP-2002) Arizona Genomics Institute, University of
Arizona, 303 Forbes, Tucson, AZ 85721, USA
The following sequence is an artificial clone that will be used for
gap filling in the rice genome. There is 60 kb of sequence from
the overlapping north clone (OSJNBa0053D03) and 60 kb of sequence
from the overlapping south clone (OSJNBa0004A10) and the gap
filling sequence begins at 60,001 and ends at 90,497 for a total
size of 150,496 bases. The gap filling sequence is 30,496 bases.
The name (OSJNAa0053D03) was derived from the location of the north
clone and the 'A' stands for Artificial. This sequence was finished
as follows unless otherwise noted: all regions were either
double-stranded or sequenced with an alternate chemistry or covered
by high quality data (i.e., phred quality >=30); an attempt was
made to resolve all sequencing problems, such as compressions and
repeats; all regions were covered by more than one plasmid
subclone; and the assembly was confirmed by restriction digest.
The following areas are below threshold:
60814, 60816, 63794-63796, 63858-63860, 68288, 68391-68392, 69501-69502,
69997, 69999-70001, 70428-70437, 70485, 70711, 71335, 71337-71339, 71392,
71398, 72046-72047, 72216-72223, 72447-72448, 73134, 18630, 80023-80024,
80097-80098, 80463, 80463, 80466-80469, 80822-80824, 80877-80881,
80918-80919, 81167-81170. The nucleotide sequence of this BAC clone
was generated by combining Monsanto, Syngenta and Arizona Genomics
Institute sequencing data.
FEATURES
source
1..150496
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/sub_species="japonica cultivar-group"
/db_xref="taxon:39947"
/chromosome="10"
/clone="OSJNAa0053D03"
/complement(608..5254)
/gene="OSJNAa0053D03.1"
/note="Contains similarity to retroelement"
complement(join(608..2290,2348..3344,3495..5254))
/gene="OSJNAa0053D03.1"
/codon_start=1
/product="Putative retroelement"
/protein_id="AAN04936.1"

gene
complement(608..5254)
/note="Hypothetical protein"
join(26554..26660,26979..27096,27264..27848)
/gene="OSJNAa0053D03.4"

CDS
19207..24030
/gene="OSJNAa0053D03.3"
/note="Contains similarity to ATP SYNTHASE GAMMA CHAIN, MITOCHONDRIAL
PRECURSOR"
/db_xref="GI:22725926"
/db_xref="GI:22725928"
/translation="MAASPAAAHHGFTITLGEAHLGRIGVGACGCEMVVPRRIAGYF
CCKQYDILLIAHFYALRTMRKSVRNKQITKAMQVAAKSLRAVQIRTESSRLQW
PFTALLGDPVSDVKKVNIIVAITSDKGLCGGINSVSKVSKALHKLTSKESKYLVI
LGEKQVQLIRDSKNIEMTVSELOKPNITQDKEMHMMWRKVVQLAEVSAVITL
QKQFTDVIATLADILKRVEDALRVFNKSHSVISFKYPTWTITLSPVMEKESSESGK
VGDLSYEIGETKSEILQNLTEFQFCVMTNACNACSELGAMSAMSSSSRNAG
EMLDRLTLTYNRTQASITTEITLIEISGALEG"
26554..27848
/gene="OSJNAa0053D03.4"
/note="Hypothetical protein"
join(26554..26660,26979..27096,27264..27848)
/gene="OSJNAa0053D03.4"

repeat_region
gene
5532..8205
/gene="OSJNAa0053D03.2"
/note="Contains similarity to flavonoid 3' - hydroxylase"
join(6532..7470,7564..8205)
/gene="OSJNAa0053D03.2"
/codon_start=1
/product="Putative chalcone flavonoid 3' - hydroxylase"
/db_xref="GI:22725927"
/translation="MDVVPLPLLLGLSVAAMVYLVYVLRGGSGGDAARRKRPLPPG
PRGWPGLNQLPOLGDKPHTMCALARYGFLRFCAEVVVAASAPVAQFLRGHD
ANFNRPNSGAHVAVNYQDLVFAPYGARWALRKLALHLFSKALDLDLRVREGE
VALVNLARQAAASVALGQANVCATNLARATIGHRVFAVDGEGEAREPEMVEL
MQLAGVNVGDFVQALWLDLPGQVAVKMKLHRRYDNNMGFINERKAGAPDGAAG
EHLNDLSVLARMQEQLDGDGKETETDILKLLNLFTAGTDTTSTTSEVWALAEI
IRHPDLKEAQHELDTVVGRGLVSESDLPRLPYLVITKFTFLRHPSTPLSPREA
RECEVDGYRIKPGATLLVNWAIARDPTQWPDLPQYQSPRELPGRMHADVDVKGADF
LIPGAGRICAGLSWGLRMTLTATLVHGFDTLANGATPDKLNMBEAYGLTLQRA
VPLMVQVPRLLPSAYGV"
19207..24030
/gene="OSJNAa0053D03.3"
/note="Contains similarity to ATP SYNTHASE GAMMA CHAIN, MITOCHONDRIAL
PRECURSOR"
/db_xref="GI:22725926"
/db_xref="GI:22725928"
/translation="MAASPAAAHHGFTITLGEAHLGRIGVGACGCEMVVPRRIAGYF
CCKQYDILLIAHFYALRTMRKSVRNKQITKAMQVAAKSLRAVQIRTESSRLQW
PFTALLGDPVSDVKKVNIIVAITSDKGLCGGINSVSKVSKALHKLTSKESKYLVI
LGEKQVQLIRDSKNIEMTVSELOKPNITQDKEMHMMWRKVVQLAEVSAVITL
QKQFTDVIATLADILKRVEDALRVFNKSHSVISFKYPTWTITLSPVMEKESSESGK
VGDLSYEIGETKSEILQNLTEFQFCVMTNACNACSELGAMSAMSSSSRNAG
EMLDRLTLTYNRTQASITTEITLIEISGALEG"
26554..27848
/gene="OSJNAa0053D03.4"
/note="Hypothetical protein"
join(26554..26660,26979..27096,27264..27848)
/gene="OSJNAa0053D03.4"

db_xref="GI:22725926"
/translation="MGSISGIDDLVFPFPQTFRRGSLDTVTNNFGKISLLSESDQSG
ENRISVFPFPNAAEYFKTISTESASNHSAEIASPTTPDQDQAGYPPVLMKLPDDL
AAVSTTTASSRRRRPKRSASTPISPSFRFVGIQPLGTSTDDQDNDVNSGTEDRRP
TDIVYDFSSRYNDDFDDFDGSLDDNVTPLYFGVMANNTEEOQARETEBERR
LEEHOAQEERLRREOQDRERAKAEDRRORALEARRADLIQODVSGTVPVFT
PQNAVAAITLLDTLLNQNELNSQSDHVUNILLQNTKMAASVPVNSVSHPTGSRVP
HFRSHDYRHSLSITGASNNRSGRHSVSPDLHRRRVERPSPRRRPIDLR
DTIIOARAARGHSDHDDDLGVAADFDDLRVDMVPAGFKTGIBSKYDGTTFES
WLTVYGLAIRAAGDSKAMANYLFLVALDSRWLHGLPRGTIGSWALKRDTFANFQ
GTPEPGTVDLVNVIQKSGESLDIRRFSEQRNKISDITDDVITIAATFGIRHEEL
VGKGRKPPRTVKLMFKSKANDHSRORSVNTPDKIMNSQCPHPNSNHAKDCFFVK
QPADQYTKARKNSDEQSISRKDDGAPAGFODHGRKELANHIFGFLAYESKRQKL
TERINAVQDPTQTLKWESETAKFKRSDHDPDRVHPGRYPLVLDTPVNVNKLRTLI
DGSALNILLFAKTLDDMQIPRSESLPSNAPPHGVIPLSATPLGQITLQVTFVGTRENF
RTNISSEVADAFETAYHAILRPAKAFMAVHYTYMMKMPGPRGVLSLRSDIKQAV
TCDKESQDMAOTREMAAREDIRLAAATASEGEIPATKTSKSGSEAKTKIPLDPS
PTKANNDKIPAWKPSMGPIPREVIEHSLHVKBDAPIKQRLRFAADRAKAIKEEL
TKLLAAGFIEVLEHPLWLANPVLVRKKTQWRMCDVYTDLAKSCPKPFGLPRIQDV
DSTAAGCELLSFLDCYSGHYRLKESDCLCTSFITPPGAYCYTWPFLGNAGATYQR
MIQFFSTQIGRNVEAYVDDVVKTKQKGLDLSLEETFSIRAFMRKLNPEKCTFGV
PSGKLGFVMSHRGIOANPEKVTAILNMKPSSTQVKQLTCMAALSFRFVSLRGEG
MPFKLLKTDNFQWQPEAKAFEDFKLLTPEPVLASPHQPELILLYVSATVSQVST
VLVVEREECHQVQRPYXVFSVLADSKTRYQVOKLYGLITLTKLGHYFQGHG
VTVVTSPLGDIHNRANGRIAKWALELMSLDSFKPRIKISQALADFAVEMTECQ
LDTPAENNEHWTMHPDGSRLSGAGVGLISPTGERLSYVLMTHFSASHNVAEYAL
LHGRIAIISLGIKRLIVRGDSQLVNVQVMKESCIDNNMAYR"
complement(608..5254)
/note="Contains similarity to retroelement"
5532..8205
/gene="OSJNAa0053D03.2"
/note="Contains similarity to flavonoid 3' - hydroxylase"
join(6532..7470,7564..8205)
/gene="OSJNAa0053D03.2"
/codon_start=1
/product="Putative chalcone flavonoid 3' - hydroxylase"
/db_xref="GI:22725927"
/translation="MDVVPLPLLLGLSVAAMVYLVYVLRGGSGGDAARRKRPLPPG
PRGWPGLNQLPOLGDKPHTMCALARYGFLRFCAEVVVAASAPVAQFLRGHD
ANFNRPNSGAHVAVNYQDLVFAPYGARWALRKLALHLFSKALDLDLRVREGE
VALVNLARQAAASVALGQANVCATNLARATIGHRVFAVDGEGEAREPEMVEL
MQLAGVNVGDFVQALWLDLPGQVAVKMKLHRRYDNNMGFINERKAGAPDGAAG
EHLNDLSVLARMQEQLDGDGKETETDILKLLNLFTAGTDTTSTTSEVWALAEI
IRHPDLKEAQHELDTVVGRGLVSESDLPRLPYLVITKFTFLRHPSTPLSPREA
RECEVDGYRIKPGATLLVNWAIARDPTQWPDLPQYQSPRELPGRMHADVDVKGADF
LIPGAGRICAGLSWGLRMTLTATLVHGFDTLANGATPDKLNMBEAYGLTLQRA
VPLMVQVPRLLPSAYGV"
19207..24030
/gene="OSJNAa0053D03.3"
/note="Contains similarity to ATP SYNTHASE GAMMA CHAIN, MITOCHONDRIAL
PRECURSOR"
/db_xref="GI:22725926"
/db_xref="GI:22725928"
/translation="MAASPAAAHHGFTITLGEAHLGRIGVGACGCEMVVPRRIAGYF
CCKQYDILLIAHFYALRTMRKSVRNKQITKAMQVAAKSLRAVQIRTESSRLQW
PFTALLGDPVSDVKKVNIIVAITSDKGLCGGINSVSKVSKALHKLTSKESKYLVI
LGEKQVQLIRDSKNIEMTVSELOKPNITQDKEMHMMWRKVVQLAEVSAVITL
QKQFTDVIATLADILKRVEDALRVFNKSHSVISFKYPTWTITLSPVMEKESSESGK
VGDLSYEIGETKSEILQNLTEFQFCVMTNACNACSELGAMSAMSSSSRNAG
EMLDRLTLTYNRTQASITTEITLIEISGALEG"
26554..27848
/gene="OSJNAa0053D03.4"
/note="Hypothetical protein"
join(26554..26660,26979..27096,27264..27848)
/gene="OSJNAa0053D03.4"
```

```
/codon_start=1
/protein_id="Hypothetical protein"
/db_xref="GI:22725929"
/tranlation="MRFTLAKVPLCRGFLIFFAHAALRHGLRLAKLITVGLIGTGY
CNRPTVSARTDGLKLLHTLLAYVAGFDVAPALSFHGSSSTNGDPCKGVDEHSYV
LGASLIDEADLTAKASARVVASQVPSAKAVVPKPKNQTVVFAFDAGLRFPCTS
LLPEILHLDFVPEFPQSPSALVRINIFDWACRTAGFEPTAEFLGALFYALVNSKVTY
PAGTKTQVFGSANFNRPERSDIPVNASPKWDRHWSKNWFLYLLMQKHEA"
29591..32802
/gene="OSJNAa0053D03.5"
/notes="Unknown protein"
join(29591..29641,29736..30492,30543..30600,30719..31208,
31289..31367,31455..31714,31790..31849,31955..32038,
32119..32469,32521..32802)
/gene="OSJNAa0053D03.5"
/codon_start=1
/product="Unknown protein"
/protein_id="AA004940.1"
/db_xref="GI:22725930"
/tranlation="MSSASPSVAPSAEYAEHLNVLATPSLSHEHHYFLGSLGNLDP
TEFIAETNRPFLRANPLNGHWKNTPKSPSLDCKTSPKSWANWTTRCVASQVWHD
ELGQALALTLANSADEPLMASAAYFWSNTILKALFNGQMPPTLLDITQITGLDV
TSSANPMSLNTQOFDRTKSIGGWSYAAVYMGKSVTPREHTAFLMLWLEKFLPCG
SSCGLTNWOLFALBETKQFPLGKLLDLYQLMSASAKIWTGSLVLSAGCESATF
DRRLPKAGADHSVEGFQDARVWPYEDSVNFDLPADFKEDINSEKPKREVFS
AALSFLPFGHOGRNIOQVSFYFHPMSARSKGWGLPIGLFFTDKIQCSGEISST
LMDLLNIPGPPILSLENIELAFRFSKFNDRWGEWKQIHFIQSASINYTDLFPDVV
PQTKSSPPRTSNRGLDIYAGLLPNGRRKTKASADPSALAPKPKTKNKKPKPA
DGLPTLDSIEQALDEIEIGEDFQAAELSDAGEKTPSPKPTPNPAAPAHFSRK
KKTAVKKSATTLKPAPQWSDRPTSATISNIEEEOPTAPAIPLVADLFSFDIK
YDFEABEETTKALAPLDEPKTKLEDISLRESLNLVDYDYSIRARFAKIOI
IPDLANITTPAVLYEQHFKLEKAKQRIADRERKDIKATIQANQLSNIDMLEARK
IDLIAQEENCAEILDEKQLANLPNAVESQKSLKSAIKNVADMTKSPKAIPTATDAQ
DIQAEEDQIRQAVSAIRYLUSA"
40478..44144
/gene="OSJNAa0053D03.6"
/notes="Hypothetical protein"
join(40478..41726,41758..41881,43886..44144)
/gene="OSJNAa0053D03.6"
/codon_start=1
/product="Hypothetical protein"
/protein_id="AA04941.1"
/db_xref="GI:22725931"
/tranlation="WADNKGKVRKSODIGTSSSRVNEAPDTSQVAVVQHLINQKLL
IEILQHRMPLNSQWQVQLEAVQALTPQVAPPTTSOKAPHAMVDPKKEASIIIC
FMCDEQHYARNCFOQKKAEMRTEDEVKMLITSTWPPPGMTKHORRSGFGAQQ
TEHLLANGRVSGSDESDQVSSDEDEKSPGFQFNQKIELKACSRCEIGHVASSCAS
TCVHCEDHPPDRCPNTKITYFFCEGTDHPKDCQFSPLTLKKWNPASSNGEKHQ
NTNPRQDRHFLTPVPGQRNRNEKRCVRREDICCFNCGQGHFADKCPKPRNIAAGT
SVNHPATCKLQAPORIVTHASRSRPIARVATIPMNALPQGVNAQFPQPPADKTTA
SICVVPLEVPILQLRNOVDREPELKKGITPTKVPNOCKIDLHITANLHEADIPVLW
SLEQTFPFGNLSVDSNDNEHEKFWNGSITALPOPSLSQPKGAPWSKSGEEIIPATG
GOAKQFRVVKFCWKEVTKQRKSKDKRRPPLNSC"
complement(47678..48706)
/gene="OSJNAa0053D03.7"
/notes="Contains similarity to retroelement"
complement(join(47678..48217,48481..48564,48635..48706))
/gene="OSJNAa0053D03.7"
/codon_start=1
/product="Putative retroelement"
```

Alignment Scores:

Pred. No.:	186	Length:	150496
Score:	9.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	20.93%	Indels:	0
DB:	8	Gaps:	0

US-10-624-218-3 (1-43) x AC131968 (1-150496)

QY 25 ArgValGlnLeuArgAsnArgVal 33

```
Db 48079 AGGGTTCAAAGGTTAAGGACAGAGTG 48053
|||||
RESULT 3
AP006237 156874 bp DNA linear PLN 31-AUG-2004
LOCUS Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 1,
DEFINITION BAC clone:OSJNBb0008D07.
ACCESSION AP006237 BA000010
VERSION AP006237.3 GI:37777305
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
REFERENCE
1 Sasaki,T., Matsumoto,T., Yamamoto,K., Sakata,K., Baba,T.,
Katayose,Y., Wu,J., Nilmura,Y., Cheng,Z., Nagamura,Y.,
Antonio,B.A., Kanamori,H., Hosokawa,S., Masukawa,M., Arikawa,K.,
Chiden,Y., Hayashi,M., Okamoto,M., Ando,T., Aoki,H., Arita,K.,
Hamada,M., Harada,C., Hijishita,S., Honda,M., Ichikawa,Y.,
Idonuma,A., Iijima,M., Ikeda,M., Ikeno,M., Ito,S., Ito,T., Ito,Y.,
Ito,Y., Iwabuchi,A., Kamiya,K., Karasawa,W., Katagiri,S.,
Kikuta,A., Kobayashi,N., Kono,I., Machita,K., Maehara,T.,
Mizuno,H., Mizubayashi,T., Mukai,Y., Nagasaki,H., Nakashima,M.,
Nakama,Y., Nakamichi,Y., Nakamura,M., Namiki,N., Negishi,M.,
Ohta,I., Ono,N., Saji,S., Sakai,K., Shibata,M., Shimokawa,T.,
Shomura,A., Song,J., Takazaki,Y., Terasawa,K., Tsuji,K., Waki,K.,
Yamagata,H., Yamane,H., Yoshiki,S., Yoshihara,R., Yukawa,K.,
Zhong,H., Iwama,H., Endo,T., Ito,H., Hahn,J.H., Kim,H.I., Eun,M.Y.,
Yano,M., Jiang,J. and Gojobori,T.
The genome sequence and structure of rice chromosome 1
Nature 420 (6913), 312-316 (2002)
22337376
MEDLINE
2247438
REFERENCE
2 (bases 1 to 156874)
Sasaki,T., Matsumoto,T. and Katayose,Y.
Direct Submission
Submitted (19-FEB-2003) Takuji Sasaki, National Institute of
Agrobiological Sciences, Rice Genome Research Program; Kannondai
2-1-2, Tsukuba, Ibaraki 305-8602, Japan
(E-mail:tsasaki@nias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,
Tel:81-298-38-7441, Fax:81-298-38-7468)
On Oct 20, 2003 this sequence version replaced gi:28564684.
Genes were predicted from the integrated results of the following:
GENSCAN1.0, BLASTN2.0, BLASTX2.0 as well as SplicePredictor
(October 1998 version). The genomic sequence was searched against
NCBI NonRedundant protein database, nr
(ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA sequence database at
RGP. Protein homologies of the coding regions were searched against
NCBI NonRedundant protein database with BLASTP2.0. ESTs represent
the identified cDNA sequences using BLASTN 2.0 with the
corresponding DBJ accession no. and RGP clone ID.
A gene with identity or significant homology to a protein is
classified based on the protein name to indicate the homology level
such as same name, 'putative-' and '-like protein'. A gene without
significant homology to any protein but with EST homology (covering
almost the entire length of partial sequence) is classified as an
'unknown' protein. A gene predicted with a gene prediction program
is classified as a 'hypothetical' protein.
The orientation of the sequence is from M13rev to -21M13 of the BAC
clone. This sequence of OSJNBb0008D07 clone has an overlap with
P0676G08 (DBJ:AP004821) clone at the 5' end and with QJ1123G09
(DBJ:AP004613) clone at the 3' end. Detailed information on
overlap and assembly quality together with annotation of this entry
is available at
http://rgp.dna.affrc.go.jp/GenomeSeq.html.
FEATURES
source
1..156874
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
```

```

/db_xref="taxon:39947"
/chromosome="1"
/clone="OSJNBb0008D07"
gene join(1483..1850,2131..2358,2408..2675,2765..3142,
3301..3435)
CDS /gene="OSJNBb0008D07.1"
join(1483..1850,2131..2358,2408..2675,2765..3142,
3301..3435)
/note="OSJNBb0008D07.1"
/note="hypothetical protein
similar to Oryza sativa chromosome 10, OSJNBa0079B05.3"
/codon_start=1
/protein_id="BAC57866.1"
/db_xref="GI:28564685"
/translation="MVELEQVVVVRTPFGSDSDASGRDPRHRRQNSPTFLRRSR
ROENPSSGSGALPPPGCGRRPRAGVRHLTYGCGSPQQAQVALLHPLVNP
PETVQWLYLVNATAORPLPATPGCTGCKAFVSLRNVPKPRPNLTKY
DGSINSEFLQIYTIIMAGGDDRVNANYPFMALKGQATFASSLTSKARTSPGKE
ADLHAVQRDDLSRSYIQFCQVRNTIPCI PAHAYIYAFRLGVHNNHLEKIASKEP
KTTTLELADKAVQEEQEBAGPLRRGLRPGCGSLRTDPAQEGHRRQQAHPPPP
PGEGRADKWCSTVHTYSLADYRTVKNLVERYKTDGERRQRREGKAPATGASDR
REEAKKAPADDDSDNLQFQSDHPTMLARGGKLALVVSPTIHNVVDGYRSVTVN
ITKTEENL"
gene join(4618..5501,5737..5792,5872..6083,6526..6620,
6681..7341,7659..8020,8148..8217,8522..8584)
CDS /gene="OSJNBb0008D07.2"
join(4618..5501,5737..5792,5872..6083,6526..6620,
6681..7341,7659..8020,8148..8217,8522..8584)
/note="OSJNBb0008D07.2"
/note="hypothetical protein
similar to Oryza sativa chromosome 10, OSJNBb0091O09"
/codon_start=1
/protein_id="BAC57867.1"
/db_xref="GI:28564686"
/translation="MIMGNSNGYRPGQCGQWQNPQPSYQCGNNGNFSNPQSLKOLV
FAQKTTDALSKMLAANDKILNTNVNQDGFASAPQNLSEFNKMIETLAQASLVIA
NESGRILQDSSSIENVAIITRGKSTDPYPNPAGTNGPKETSDADKEIQP
EKIVQEYVLTLLPQSRKPLVDEQFAFVEIQIHISIPLLDAMQVPTVAHNL
KDIILNRPLPTTEVLLPQTHCSNLIHLKPEKKDPCPTITCSIRAQDFQALCDL
GASISVMPKLSLTSSTRKSLSPNQGNNAPSSSEKLPKNGNNAQELLEDAKSI
TCTQEVRAVSSEAVSRTIEKEGVEEAKDARSITSGDGWKIRELNKEFALKFIHA
LIVFRLSKLSVAYIQVALVLTGIRRLRLFLFGSGSPSKRHSDESTRSDVSMED
ADAPQLVDDDLVDLSRQAYMLSDWEYAHMREYSPPELLKKISMDVFRSLWKA
WGMQFAVDPGSLDLTLQFLCTKEVEDGISFCFKEKFTLTWKSLSLTLGHDSY
KIDLLKGIITFAMTIEESRAGTYRETRNMTNEQSSSSSTPMQWAGWAPIGDGP
IRGDELITTEAMTIEESRAGTYRETRNMTNEQSSSSSTPMQWAGWAPIGDGP
WTQAPRHIGVSTWSTSDERHAPHDIHWGNQPSRSRGVPPAPSEHSSSRWDLG
EITRRMDTLDMQTGEIRGSIFIEDPERKPSLGGDIATPTEVVARYEIVNMSGYDPTDS
LA"
gene complement(join(9192..9253,10014..10099,10349..10648,
10691..10753,12873..12884,15000..15124,15196..15408,
16077..16148))
CDS /gene="OSJNBb0008D07.3"
complement(join(9192..9253,10014..10099,10349..10648,
10691..10753,12873..12884,15000..15124,15196..15408,
16077..16148))
/note="OSJNBb0008D07.3"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAC57868.1"
/db_xref="GI:28564687"
/translation="WVFSRWQAWYPSNDKLSIAHNKWLAPSSLSAVVAPDLLPLRH
RLRLRLVLAIVPFPADRNAAVLVDPTRAAASLRFGRFRFRFRPGVAAQDHR
GAGAPPGVGCAGEAAGRAPRLPCGARSAAALGNQPKTKNRLVDYDDIAGHGS
EPTPMWRGRCELTSVKNNDGRNRTKAGDELRLGTRQRRRSGLRLKRGNGSA
PTTFREYGDGAXRRLRPPQRRGAARTTTTAAARTRWRRYQRDQKAGFGGRRRVA
SVRQPEVEDQEGGLYVELDKQDTWHTCIIYH"
gene complement(join(17930..18993,19111..19156,20461..20505))
CDS /gene="OSJNBb0008D07.4"
complement(join(17930..18993,19111..19156,20461..20505))
/note="OSJNBb0008D07.4"
/note="hypothetical protein
similar to Oryza sativa chromosome 10, OSJNBb0036B06.13"

```

```

/codon_start=1
/protein_id="BAC57869.1"
/db_xref="GI:28564688"
/translation="MGWKLCHFRPYTVMPNREKKNVSGYASNQNYVDFRPAFQ
VDELPMPEEDVPFCHDKREOCHVALNCPWYVYAKEAPSGLAGQASLNQPPRTKSLQ
VSQSRPENGPNLPPRCFPILSYHVGSRTKQAKGSKKNQKGAKNRTPBEEVKMT
ITSPEMPARTITIPRQRGNTFRGAROLTOHLLSHGGRVSDSGSDVSSDDDESPR
HLGEEMESKACTRCGEIGHVASMCLILCPHCEEDHPGECPTKVTGCLCEGTNHLV
KDCOLSLLTKTAIEIQRITLRLSHOLVSSNSINIDNIITHNTNAPPAPPSAICGNSNTF
STNHSLRHQPOQRIIVDGRSCVSPRPTPGRAKGNOSQEGDTKQKTVLCFK"
gene 21714..22220
CDS /gene="OSJNBb0008D07.5"
join(21714..22220)
/note="OSJNBb0008D07.5"
/note="hypothetical protein
similar to Oryza sativa chromosome 1, P0489B06.21"
/codon_start=1
/protein_id="BAC57870.1"
/db_xref="GI:28564689"
/translation="MAAAGRARRRPRHRSRHEATPRPDGRRRRPPARRSGARBERGGK
GEAKWRSAPGSGTATTETALEARGGAARVNEDSGVPVAGGRNEGVEGVEDAAKP
KEAMPKREETRGDDGGPKLNGDGERRRLSDGGERERRAAETEEGGRWABEAPQE
GKGGKRTF"
gene join(23847..24020,24212..24301)
CDS /gene="OSJNBb0008D07.6"
join(23847..24020,24212..24301)
/note="OSJNBb0008D07.6"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAC57871.1"
/db_xref="GI:28564690"
/translation="MRRMEDEGGGKKRQRREMDAPLRLSSRPSLGRGCLRALQW
LDVEGRGRGRGRELWRYVNEARETGTGELGRNRRMESENSA"
gene join(27649..27903,27964..28167,28221..28568)
CDS /gene="OSJNBb0008D07.7"
join(27649..27903,27964..28167,28221..28568)
/note="OSJNBb0008D07.7"
/note="hypothetical protein
similar to Oryza sativa chromosome 10, OSJNBa0034B05.19"
/codon_start=1
/protein_id="BAC57872.1"
/db_xref="GI:28564691"
/translation="MKRVLTGAGGGGHWASLNIISLAAPHALKAPQMRLOPSLPIGV
TPGHTWPIGHELLVIFNSNTFRTEWIDFVANLNLPTNTHKPGNPGPIVTSQDVK
ILAAACQADSLPAATPEEPAGGAPASASRASKRITSGDEVFPVKETPLPSVLSVF
LRANSNVFAWPSDMPGVPREVIEHRLAMPDTPRIQKVRQAPERQAYIRREVTRL
LEAGTIREVHPVWLVNPAVVPKANGKLRMCIDYTDLNKACPKDPSLYHT"
gene 29246..29641
CDS /gene="OSJNBb0008D07.8"
join(29246..29641)
/note="OSJNBb0008D07.8"
/note="hypothetical protein
similar to Oryza sativa chromosome 1, B1064G04.25"
/codon_start=1
/protein_id="BAC57873.1"
/db_xref="GI:28564692"
/translation="MGIKLCFASPANPKSNQGVHANAEIVKGLKTRTVNVLKKGDS
WLEELPAVLWANRTTPSHATGTPTFVLYGAELVFLPSLSLGSFRVALYNANQDELRL
RDDLDYLEERRRRRAALWAARYQQLSRCYH"

```

Alignment Scores:

Pred. No.:	193	Length:	156874
Score:	9.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	20.93%	Indels:	0
DB:	8	Gaps:	0

US-10-624-218-3 (1-43) x AP06237 (1-156874)

Qy 21 GluLeuArgGlnArgValGlnArgLeu 29

Db 139455 GAACCTTAGACGCGAGTTCAGAGACTT 139481

Query Match:	20.93%	Indels:	0
DB:	8	Gaps:	0
US-10-624-218-3 (1-43) x AC137747 (1-175461)			
QY	24	GlnArgValGlnArgLeuArgAsnArg	32
DB	104259	CAAGGGTTCAAGATTGAGGACAGG	104285
RESULT 5			
AC021892	199113 bp	DNA	linear
LOCUS	Genomic sequence for Oryza sativa, Nipponbare strain, clone		
DEFINITION	OSJNBa0053D03, from chromosome 10, complete sequence.		
ACCESSION	AC021892		
VERSION	AC021892.16	GI:18254440	
KEYWORDS	HTG.		
SOURCE	Oryza sativa (japonica cultivar-group)		
ORGANISM	Oryza sativa (japonica cultivar-group)		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.		
AUTHORS	1 (bases 1 to 199113) McCombie, W.R., Cordum, H., Minx, P., de la Bastide, M., Spiegel, L., Nascento, L., Zutavern, T., Ballija, V., Bell, M., Baker, J., Santos, L., Miller, B., Katzenberger, F., Muller, S., Preston, R., Kirchoff, K., Kuit, K., King, L., Yang, C., O'Sahughnessy, A., Palmer, L., Dedhia, N. and Wilson, R.		
TITLE	Genomic sequence for Oryza sativa, Nipponbare strain, clone		
JOURNAL	OSJNBa0053D03, from chromosome 10, complete sequence		
REFERENCE	2 (bases 1 to 199113)		
AUTHORS	McCombie, W.R.		
TITLE	Direct Submission		
JOURNAL	Submitted (22-JAN-2000) Lita Annenberg Hazen Genome Center, Cold Spring Harbor Laboratories, 1, Bungtown Road, Cold Spring Harbor, NY 11724, USA		
REFERENCE	3 (bases 1 to 199113)		
AUTHORS	McCombie, W.R.		
TITLE	Direct Submission		
JOURNAL	Submitted (22-JAN-2002) Lita Annenberg Hazen Genome Center, Cold Spring Harbor Laboratories, 1, Bungtown Road, Cold Spring Harbor, NY 11724, USA		
REFERENCE	4 (bases 1 to 199113)		
AUTHORS	Palmer, L.E., Yu, M., Cordum, H., Minx, P., de la Bastide, M., Spiegel, L., Nascento, L., Zutavern, T., Ballija, V., Bell, M., Baker, J., Santos, L., Miller, B., Katzenberger, F., Muller, S., Preston, R., Kirchoff, K., Kuit, K., King, L., Yang, C., O'Sahughnessy, A., Dedhia, N., Wilson, R. and McCombie, W.R.		
TITLE	Direct Submission		
JOURNAL	Submitted (02-APR-2002) Lita Annenberg Hazen Genome Center, Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring Harbor, NY 11724, USA		
REMARK	Genomic sequence for Oryza sativa, Nipponbare strain, clone		
COMMENT	OSJNBa0053D03, from chromosome 10 On Jan 22, 2002 this sequence version replaced gi:14209718. This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest.		
FEATURES	Location/Qualifiers		
source	1. .199113 /organism="Oryza sativa (japonica cultivar-group)" /mol_type="genomic DNA" /cultivar="Nipponbare" /db_xref="taxon:39947" /chromosome="10" /clone="OSJNBa0053D03" complement(<1. .4300)		
gene			

AC137747	175461 bp	DNA	linear	PLN 28-JUN-2004
LOCUS	Oryza sativa (japonica cultivar-group) chromosome 5 clone			
DEFINITION	OSJNBa0051L16, complete sequence.			
ACCESSION	AC137747			
VERSION	AC137747.2	GI:49258209		
KEYWORDS	HTG.			
SOURCE	Oryza sativa (japonica cultivar-group)			
ORGANISM	Oryza sativa (japonica cultivar-group)			
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.			
AUTHORS	1 (bases 1 to 175461) Chow, T.-Y., Hsing, Y.-I. C., Chen, C.-S., Chen, H.-H., Liu, S.-M., Chao, Y.-T., Chang, S.-J., Chen, H.-C., Chen, S.-K., Chen, T.-R., Chen, Y.-L., Cheng, C.-H., Chung, C.-I., Han, S.-Y., Hsiao, S.-H., Hsiung, J.-N., Hsu, C.-H., Huang, J.-J., Kau, P.-I., Lee, M.-C., Liu, H.-L., Li, Y.-F., Lin, S.-J., Lin, Y.-C., Wu, S.-W., Yu, C.-Y., Yu, S.-W., Wu, H.-P., Shaw, J.-F., *Yu, Y., *Rambo, T., *Currie, J., *Collura, K., *Soderlund, C. and *Wing, R.			
TITLE	Oryza sativa BAC OSJNBa0051L16 genomic sequence			
JOURNAL	Unpublished			
REMARK	*Arizona Genomics Institute, University of Arizona, 303 Forbes, Tucson, AZ 85721, USA			
REFERENCE	2 (bases 1 to 175461)			
AUTHORS	Chow, T.-Y. and Hsing, Y.-I. C.			
TITLE	Direct Submission			
JOURNAL	Submitted (29-NOV-2002) Institute of Botany, Academia Sinica, 128, Section 2, Academia Road, Nankang, Taipei 11529, Taiwan			
REFERENCE	3 (bases 1 to 175461)			
AUTHORS	Chow, T.-Y.			
TITLE	Direct Submission			
JOURNAL	Submitted (26-JUN-2004) Institute of Botany, Academia Sinica, 128, Section 2, Academia Road, Nankang, Taipei 11529, Taiwan			
REFERENCE	4 (bases 1 to 175461)			
AUTHORS	Chow, T.-Y.			
TITLE	Direct Submission			
JOURNAL	Submitted (28-JUN-2004) Institute of Botany, Academia Sinica, 128, Section 2, Academia Road, Nankang, Taipei 11529, Taiwan			
COMMENT	On Jun 26, 2004 this sequence version replaced gi:25901064. This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >=30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by more than one plasmid subclone; and the assembly was confirmed by restriction digest. There are Bacterial Transposons at the following locations: 41758-45686, 47895-52103, 78702-81437, 101398-103971, 128754-131079, 133326-135977, 140916-144022, 157617-159911 and 164478-166286. The areas listed are covered by bacterial transposon reads only: 80517-80599, 134314-135151, 158162-158858 and 165339-165423. There is a single subclone are from 139434-139525 and an area only covered by PCR at 1288-1310. The nucleotide sequence of this BAC clone was completed to Phase 2 by the Academia Sinica Plant Genome Center and Finished/Annotated (Phase 3) by the Arizona Genomics Institute.			
FEATURES	Location/Qualifiers			
source	1. .175461 /organism="Oryza sativa (japonica cultivar-group)" /mol_type="genomic DNA" /cultivar="Nipponbare" /db_xref="taxon:39947" /chromosome="5" /clone="OSJNBa0051L16"			
ORIGIN				
Alignment Scores:	212	Length:	175461	
Pred. No.:	9.00	Matches:	9	
Score:	100.00%	Conservative:	0	
Percent Similarity:	100.00%	Mismatches:	0	
Best Local Similarity:	100.00%			

CDS	repeat_region	/gene="OSJNBa0053D03.1" /note="Putative retroelement" complement(join<1. .2305,2518. .3323,4256. .4300)) /gene="OSJNBa0053D03.1" /note="Similar to RNA-DIRECTED DNA POLYMERASE HOMOLOG (REVERSE TRANSCRIPTASE HOMOLOG) E-value=4e-39 Identites=81/142 (of subject); Contains pfam domain rvt Reverse transcriptase (RNA-dependent D e-value=1.5e-45; Contains pfam domain zf-CCHC Zinc knuckle e-value=6.2e-05" /codon_start=1 /product="Putative retroelement" /protein_id="AAM00937.1" /db_xref="GI:19881536"	
		/translation="MGWKLCHVFRPYTVVMAGSVTRRLGSMTGFLVLELGMVVLGYG HGVDYQARLPQEGDEAGHAGVEVAVILSGSPERTSLAVTAGGDSFPACQNAAL LAICTHORVYDELOHSPYRYPHRRGGARDVATFRDASSEDATVHLARVVEAYDA RIDHQMVRGMVNLKILRLQENLOKLDLAEQLHQLKIAOGESRPRKRRV CRSQITARKSTPRLVSGDYDFQRLVDSIKSIRLEAKHELESKRLANFRNOOGAN FRPGFSRSLSFAGSVGMFQVTVTRNVNTGEGNQEGSNHNGQNPPTPPPPPP PPDTNALITQILAQANMTAFLLHLQNPQONAPPQPPQHSKLAELRIRPPTFSS NNVDALDMLHAGVKLDTQCSDEEKVFAAHLQGLQSPASLWMDHFOATQPEGOPITW ARFTAAPRTHVAGVVALKKREFELKQGNRSVMEYLHEFNMLARVAPQDVRDEEK OEKLAGMDPELSVRLVSGDYDFQRLVDSIKSIRLEAKHELESKRLANFRNOOGAN QVRVTNPFGSSSQQQQQQPPSARPPQFVRIIPQPPQQQQQCTRAARPPTPTV QPVGDRDAQOQLRCFCFEPGFADCKPKPRQQQAPPNSNGKDVIRGRNVH TAEDVLTTPDVIQVFLVHSIPATILFDSGASHFISVSVFGRNQLGVRLRNPLLIT TPGVMITAKYISPAVPIBIQIPIPPSLDILDLTKNDLVILGNMLAQFGVVDCAERT VTLVRGPQPVVFAPPTSVRSSELHQLGLSEIPIVREFGVDFPEELPGMPPKREIEF RIDLAPGTTVLHKPYRMANELAEVKQLBEKGVIRPSTSPMGAPMIFVKKDK TKRMCVDVRLNEVFNKPYLPRIIDDLDFQDKGATVFSKIDLSRGYHQLIRREDIP KTAFTRYGLVEFTVMSGLTNAPAFNMNMKNVMEYLDKRVVVFIDDIILYIYQSE DHQHLRLVLGKLEHQLYAKL" complement(1. .3507) /note="Similar to putative retrotransposon, Ty3-gypsy-like" complement(1391. .1414) /note="Simple repeat (CAG)n" complement(3509. .5087) /note="Similar to putative retrotransposon, Ty3-gypsy-like" 4880. .4908 /note="Simple repeat (TAA)n" complement(5448. .6661) /note="Similar to putative retrotransposon, Ty3-gypsy-like" 5583. .5584 /gene="OSJNBa0053D03.2" /note="Hypothetical protein" 5583. .5984 /gene="OSJNBa0053D03.2" /note="Similar to (AC084762) putative splicing coactivator subunit [Oryza sativa] E-value=2e-09 Identites=43/133 (of subject)" /codon_start=1 /product="Hypothetical protein" /protein_id="AAM00938.1" /db_xref="GI:19881537" /translation="MVTAGDHRGGAAPERAERCKANGGLRLTPQRRRRRRPERR KAARFGMTMTVLRSGAERRTGSAATRSRSGWRCRVGRGTATTAARSSAATA EREHGASTDPTARASGERRRKKSARGAIL" complement(6696. .7683) /note="Similar to putative retrotransposon, Ty3-gypsy-like" complement(7800. .7876) /note="Similar to 5S rRNA gene-like" 7882. .7961 /note="Similar to Rice transposon Tnr3 DNA, complete sequence" complement(8199. .8275) /note="Similar to 5S rRNA gene-like" 8281. .8363 /note="Similar to Rice transposon Tnr3 DNA, complete sequence"	
CDS	repeat_region	8363. .10391 /note="Similar to putative retrotransposon, Ty3-gypsy-like" 9254. .12323 /gene="OSJNBa0053D03.3" /note="Hypothetical protein" join(9254. .9369,9374. .9689,9856. .9985,10451. .10797, 10984. .11108,11163. .11272,11357. .11520,11826. .12323) /gene="OSJNBa0053D03.3" /note="Similar to (AC022352) Putative retroelement [Oryza sativa] E-value=e-105 Identites=223/1175" /codon_start=1 /product="Hypothetical protein" /protein_id="AAM00939.1" /db_xref="GI:19881538" /translation="MTSAMTSAMTSAPAAARLSTCSSANDGAKARTETNGNGRRGE LTQNGGSGRTDGDGVEGEAKIGSTAAATVLRQSSATAKERTMATTPRRWRPSR TGNGKDPNLEGGDGAARLGPRAAAGASFTSGPRGVTVPASVCSQNFQFPICQOTQ GAPQORADMIADVMEQFGKLPKDTGNLYRHYPPEWFERVPLPARYKVNFSESG QDNVSTYEHASFLAQCEASVDALRDYIQFRRETRNRCYSLSALTDLSQLADLAFQGL IAPITKEFSSQFAEAKNFKKINHVIYMYGSDDEEDSVAAMVRSKRYCKWHNS GSKYNTCKVFRQOIQVAIEGGKIKFDDSKPKMGVDGNPFPVNMIGPNQDCQDDDE DRYKADAKRPAKGSVVATSNLVFTLTAEFGAEQADVDEVEASARLVLSLEQAVF EKPGTDNRHLKPLFYNGYVIGKPKSMKMDVGLL" complement(10588. .11006) /note="Similar to putative retrotransposon, Ty3-gypsy-like" 12566. .15812 /note="We believe the assembly to be correct. There are six copies of a closely related repeat that are arranged in three groups of two (to be designated as Repeat A, Repeat B and Repeat C). These copies have been sorted based on forward-reverse pairs and base pair differences. The assembly is consistent with the relevant fragments of the HindII digest. RepeatA: Beginning of First copy of RepeatA (A-1) 12566 End of the First copy of RepeatA (A-1) 15812" complement(12568. .14225) /note="Similar to putative retrotransposon, Ty3-gypsy-like" 14024. .14050 /note="Simple repeat (TAA)n" complement(14589. .15811) /note="Similar to putative retrotransposon, Ty3-gypsy-like" complement(14811. .15188) /gene="OSJNBa0053D03.4" /note="Hypothetical protein" complement(14811. .15188) /gene="OSJNBa0053D03.4" /note="Similar to (AP002867) hypothetical protein [Oryza sativa] dbj BAB17324.1 (AP002747) hypothetical protein [Oryza sativa] E-value=5e-21 Identites=57/290 (of subject)" /codon_start=1 /product="Hypothetical protein" /protein_id="AAM00940.1" /db_xref="GI:19881539"	
		/translation="MEPRSPRIPISPAALSPSRCYIIVPRASLFRFRSRPLALAVGF VLAPSSLSLSDVAALRFTVAAPDFLPTRRHLRURVVVDLARSFVSPADRRSTVAL VIPNRAAFLRSVRRLRRPRVSG" complement(18090. .21141)	
gene	repeat_region	235 9.00 100.00% 100.00% 20.93% 8	
		Length: 199113 Matches: 9 Conservative: 0 Mismatch: 0 Indels: 0 Gaps: 0	
gene	repeat_region	235 9.00 100.00% 100.00% 20.93% 8	
		Alignment Scores: Pred. No.: 9 Score: 9 Percent Similarity: 100.00% Best Local Similarity: 100.00% Query Match: 20.93% DB: 8	

Alignment Scores: 259 Length: 223255
Pred. No.: 9.00 Matches: 9
Score: 100.00% Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 20.93% Gaps: 0
DB:

US-10-624-218-3 (1-43) x AC132702 (1-223255)

Qy 11 GluSantHrAlaLeuGluThrArgVal 19
|||||
Db 215309 GAAACACTGCGTGGAGACCGAGTT 215283
|||||

RESULT 7

LOCUS AEO17294 300207 bp DNA linear BCT 22-MAR-2004
DEFINITION Leptospira interrogans serovar Copenhageni str. Fiocruz LI-130,
chromosome I, section 8 of 14 of the complete sequence.
ACCESSION AEO17294 AEO16823
VERSION AEO17294.1 GI:45600841

KEYWORDS

SOURCE Leptospira interrogans serovar Copenhageni str. Fiocruz LI-130
ORGANISM Leptospira interrogans serovar Copenhageni str. Fiocruz LI-130
REFERENCE 1 Bacteria; Spirochaetales; Spirochaetales; Leptospiraceae; Leptospira.
1 (bases 1 to 300207)

AUTHORS

Nascimento, A.L., Ko, A.I., Martins, E.A., Monteiro-Vitorello, C.B.,
Ho, P.L., Haake, D.A., Verjovski-Almeida, S., Hartskeerl, R.A.,
Marques, M.V., Oliveira, M.C., Menck, C.F., Leite, L.C., Carrer, H.,
Coutinho, L.L., Degraeve, W.M., Dellagostin, O.A., El-Dorry, H.,
Ferre, E.S., Ferro, M.I., Furlan, L.R., Gamberini, M., Gigliotti, E.A.,
Goes-Neto, A., Goldman, G.H., Goldman, M.H., Harakava, R.,
Jeronimo, S.M., Junqueira-De-Azevedo, I.L., Kimura, E.T.,
Kuramae, E.E., Lemos, E.G., Lemos, M.V., Marino, C.L., Nunes, L.R., De
Oliveira, R.C., Pereira, G.G., Reis, M.S., Schriefer, A.,
Siqueira, W.J., Sommer, P., Tsai, S.M., Simpson, A.J., Ferro, J.A.,
Camargo, L.E., Kitajima, J.P., Setubal, J.C. and Van Sluys, M.A.
Comparative Genomics of Two Leptospira interrogans Serovars Reveals

TITLE

Novel Insights into Physiology and Pathogenesis
J. Bacteriol. 186 (7), 2164-2172 (2004)
15028702

JOURNAL

PUBLISHED 2 (bases 1 to 300207)

REFERENCE

Nascimento, A.L., Ko, A.I., Martins, E.A., Monteiro-Vitorello, C.B.,
Ho, P.L., Haake, D.A., Verjovski-Almeida, S., Hartskeerl, R.A.,
Marques, M.V., Oliveira, M.C., Menck, C.F., Leite, L.C., Carrer, H.,
Coutinho, L.L., Degraeve, W.M., Dellagostin, O.A., El-Dorry, H.,
Ferre, E.S., Ferro, M.I., Furlan, L.R., Gamberini, M., Gigliotti, E.A.,
Goes-Neto, A., Goldman, G.H., Goldman, M.H., Harakava, R.,
Jeronimo, S.M., Junqueira-De-Azevedo, I.L., Kimura, E.T.,
Kuramae, E.E., Lemos, E.G., Lemos, M.V., Marino, C.L., Nunes, L.R., De
Oliveira, R.C., Pereira, G.G., Reis, M.S., Schriefer, A.,
Siqueira, W.J., Sommer, P., Tsai, S.M., Simpson, A.J., Ferro, J.A.,
Camargo, L.E., Kitajima, J.P., Setubal, J.C. and Van Sluys, M.A.
Direct Submission

TITLE

Submitted (29-FEB-2004) Laboratorio de Bioinformatica/Instituto de
Computacao, Universidade Estadual de Campinas, Avenida Albert
Einstein 1251 Box 6176, Campinas, SP 13084-971, Brasil

FEATURES

source Location/Qualifiers
1..300207
/organism="Leptospira interrogans serovar Copenhageni str.
Fiocruz LI-130"
/mol_type="genomic DNA"
/strain="Fiocruz LI-130"
/serovar="copenhageni"
/db_xref="taxon:267671"
/chromosome="I"
complement(274..1461)
/locus_tag="LIC11736"
complement(274..1461)
/locus_tag="LIC11736"
/note="Identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"

gene

CDS

/codon_start=1
/transl_table=1
/product="conserved hypothetical protein"
/protein_id="AAS70325.1"
/db_xref="GI:45600842"
/translation="MFIQFFVRLKAEGIPVTTGFLDFLKFIDYTTTHKKAWIGLNEL
YPSRACMVNDIKYFDADFVSEIFGREGALKEPDLQELFNWLNELFDNKLPPSL
IPDQLLEELKRLBEQGHGKNKVGCGSGPGHGGKNGPVGRIIDGAGNSAI
FQALRRYREYTDQIDIRQIKTAKRVKRLKEGVSEFHLKPIITDSTCKNAGDIQV
EFERSKNGKLVLLMDTGGSTMSHAERVKNLFSASHQINHFKEFYHFIYFNIIYDAI
YPRANMRTPISLQRMFKHSEDITKLIDGAYMAPYELLDSTGYGYSRFRMFLRPE
NPESGLDSLKKIRRHFRDSIWNPEPIKYWDAPTQIEIGNQVPMVFLSLDGLGKIKK
LIN"
complement(1463..2260)
/locus_tag="LIC11737"
complement(1463..2260)
/locus_tag="LIC11737"
/note="Identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"
/codon_start=1
/transl_table=1
/product="conserved hypothetical protein"
/protein_id="AAS70326.1"
/db_xref="GI:45600843"
/translation="MKLSTETLLSPALEAIIAENVTSRPLLKGEPTGKSLAIEY
LADQKPLTYWHIKSIQAKGEGFYFDVSRNLSRSESESKVKNENIIRLGAIG
EAFQNKQSLVIDEIDKADIEFDNLELDORMEFYIPEISKRIQAOHRIITLISN
NEKELPAFLRRCIHFYIEFPDPEMKIIVSHFPGIGHTLLIKALEMFLYLRNMDLL
KKRPTSELDDWIOILVHQGAVLKKEVRIPPLGALIKNEEDLRFPN"
complement(2313..4442)
/gene="pbbp2"
/locus_tag="LIC11738"
complement(2313..4442)
/gene="pbbp2"
/locus_tag="LIC11738"
/note="Identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"
/codon_start=1
/transl_table=1
/product="penicillin binding protein"
/protein_id="AAS70327.1"
/db_xref="GI:45600844"
/translation="MIQRTFKIRIMPKMIRLILWLLSIFLYTNLFPSEKNTETIPS
YKEIRNSRPSDGVLDHGHRLQIIRNVRERKLSWTEGEIIPETFLALFLQEDKR
PFHSGVDRIALGSIKDRFNRRGASTLSMQLAGFLTKGQQRNIFQKWEQMT
AQIEKTWTNKEILTAVNLTPFRGELGRASRGLFQKEPSTLSDTESILLVAMP
YPGASYVLAKRSICILAKIOKEELCDYFESVAKKATSKNNLPSTEGIAVHAKIF
RENPEISIDGKINTIDFDLQWKITETAKNNLYCLKKQNVSETGILVDNISGAVLA
YIGNLEDSNIFYDAIQSKRQAGSTLKPFLGLAFKEILKPNLSLESPAEWNAVSG
IYKPSNYSDTYHGVNQAKYALASNLIPAIRVLDLVNVPDFVDLKLGLGLKGRADP
YKSGALGTADTVTLFELTNAYRTLANGISISKPTFFPEAKRTVOENKFNWNRVY
TKSGADTLSEILSDREYRSLSGLNNYLSLSTRTFTAVKTGTQDMRDNKICIGSKYTV
GWVYNNGKPMWDVSGVTGAAPTNTVTINLQREEOSNNLTKNYSIYPPAROLTEL
SKIPKILLPGNETIYALDPPDPEGRQKLHFASQPDGFKWLDGSKSQEAKKEVFVW
KPEKGFHLISLQDQNGKSIDSVVFVR"
complement(4606..10410)
/locus_tag="LIC11739"
complement(4606..10410)
/locus_tag="LIC11739"
/note="Identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"
/codon_start=1
/transl_table=1
/product="conserved hypothetical protein"
/protein_id="AAS70328.1"
/db_xref="GI:45600845"
/translation="MNHUFGQTLKILLNSFTILSRSSISFYLKLSIKIDSLKNFY
FFQKVNFFIYINWLVKXNAFVCGSLFFFAFSTLNAQIKIEPSTGEVKKPSQIRA
RFSVSMPLGNPKFSLPFEIRCEPLQGTQRWVDKNWLEFPELLPGGIECTFETKKI
KSVAGNPLNKEGKTSFFFGPEIENHTFPYGVYVDEDDQIFILNLDIDIRSSANDH
IYFVVEGLDKDIGSYKVSVEKEILKVRTIKTQKTDRLMIKPDQKPSGQKVLVLEK
GLKSGVPSRSTRIEYSVQAFAEFSRCDRVNAKACIPLSLPLAFNSPVSEIL
KKIQLTNDGKTIPIVKVFSENGHDQYDVSPVLPSPKSKFQILLPLPGIKDDSGRTLSN

QSSPFLTVTDDYPILLKASTFTGILEREPEALVPTIENLEAENPVRLYQVKTSPDT
QDKLQFOLKKEKEKELINWAIKKEKESKNSPPKQFTKGLILSGEIAELRLVLT
IENLDSNGIFDPSQNPOTNDIKLQSNHGARFEVVGILPKKSGHVVADSDVLGN
SLIINGQPFVTRTAVLWNLALFKWGSSESLVWTKNDSPKPNADIOIFNCKNEK
IFTGKPGSTLTITKGLSKNKTIPHCWSKSYENGLFTASFEDDFTFTHQWNGIEN
WRNLPSGDVGNENVVPHPLDRTLFRAGEIASIKLVRSIKKSFGEIPSQNEYPFAFA
KLVHSGNSKEYSISLAKWNGEGTSSFOFKIPKANLCTYOILLPGEGSIGPEVDFRV
PLMKAEIQSGNSINERKSGITGSVRYLSGGGAGKLPVLLTRVTLSESGAYPDYSEF
SFNGKVNQBEISEENTFTKTLILDERGFESTVRISIPSDTTQKLETELEYRDP
NGEIQSVRYFPFTYPSQYHIGITSEGAAVQVDSKLVKAVLDLKGPIISGKKVNTAF
TKKYSNRKLGVGFYSYEHKSEVKGELGFCSGKTSKGLFYNGSLKTKTYIYFAS
LSGEBIKANSWMTIGRDDIWFASDHRMDLIPKKEYQTGEGKAKFQVMPREFATA
LVTEVERGILNSITGNLSGKEPVIIEVPIESSFAPNVSFVSLAVRGVDPGKETAVDL
AKFSRLGMAQIRVHKPFEPVPIETDKTVGTROKAKVKIQLIDHPSIOVKNKSGG
LAAVDGLLEKNNWDLKAMNORGNSVQTSIAQLHVGRHRHGLSLPFGGGIT
GATTELDLFLKPDLPKDPENGLEVEIPNDLSITSFKIVAIVHSGHVDKSGSGTQ
VAILVSVKDPNKEDLPRLIRGAISROQKQHMDDITPANALGILAFGLWLLSSNDST
VEGTVTVLNNSTLEWKNQKPNKLTLPMPHNAQNLFEVQNGKGPVYVHTKAL
PLKEKLSGRMLEKILNESGNKTSFOEGDIVRVLKIVTESDLSWIAVRDPIPAAG
SILSGCLGNSRSGSELTKENWMSPTETIERKWEVTAIFYEVLPGSVTLVYVIRIN
QTGKPLPPTKRVEMVILPDQFAELPNSDQMITKE"
complement (10484. .11818)
/gene="pmba"
/locus_tag="LIC11740"
complement (10484. .11818)
/gene="pmba"
/locus_tag="LIC11740"
/note="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer; Pmba"
/codon_start=1
/transl_table=11
/product="peptide maturation protein"
/protein_id="AAS70329.1"
/db_xref="GI:45600846"
/translation="MNLEHSVYVLDVCKKKGDLGDFDLVASESKDVIGIELFRKRVNST
ELNSRGIGIRLIRSGKPGYSYKSEALSQMVDEAISQKISDPLDIDLPGPTSL
PKIRSYEEGLSEGLFWLSTGKLDLWASVGGKIENVPYSYAGTKWSTIRILANS
NDLVHSEKNSLVAGVALVADTGTKMMGYTRSGDLKDLPINELIVSTATERSLALL
CAKPKVSGSYILLNRLSPQIFGMFSSPFSADSVQKLSRLDKGVQSQASQNVNVP
CDPHSIDPGRLLDAEVLTRKKAIIENGILKTYLNLSSAKAGVSPGTGNVRSYG
GRVTSFNYPVKNQTLLELLKSVPECIYILKLEGSGCSAVSGISIGVQGIYYK
NGNPVHPVDNITMNLNPFLLFRIDGISNEYNDSSYSIKIPDILLIREASIAG"
complement (11815. .13200)
/gene="tldb"
/locus_tag="LIC11741"
complement (11815. .13200)
/gene="tldb"
/locus_tag="LIC11741"
/note="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"
/codon_start=1
/transl_table=11
/product="tldb"
/protein_id="AAS70330.1"
/db_xref="GI:45600847"
/translation="MNINKAQLILEAGLSRKADFVEIPEEBTRSSSVSLDRDKIEQSF
AGIDYGIRLYGTDLVYAHTNEDLEHLSLIDLADRSATKESQRTVLKGLDKL
IPSFSTNLIDPRKISPDQKDLILYKADTARNVSNIVQVGVSAFDSVSRIGIYNSEJ
LSLEDLVRKRSFINVTAKEGGERFVASEACHPLEATEIRKKSSEFPVDRKGRKIAQSCL
MLSAGIIEGKNVGMNGFVGGIIRKKSSEFPVDRKGRKIAQSCL
TADIDGTIPDSNGSINVDDEGAPQKTLIENGILKNYLSDRVGAQVEVGEKTSARR
ESYAVPYSRMNTYIAAGKDSFDSMLSGIDIGLFAKTKWGGGVSNPSTGEFNFVSSE
VYIRNGKIAEPRVGATLICKGDEILPKISWVGKOLELAAGCGAASGSVPVTVGQPSL
KVEILVGGRS"
complement (13249. .13626)

Alignment Scores:

Pred. No.: 331 Length: 300207
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 20.93% Indels: 0
DB: 1 Gaps: 0
US-10-624-218-3 (1-43) x AE017294 (1-300207)
QY 19 ValaLaGluLeuAtcGlnArgValGln 27
|||||
Db 227460 GTAGCGGAGCTAAGACAAAGAGTTTCG 227434
RESULT 8
LOCUS AE017074/c 304110 bp DNA linear PLN 06-JUN-2003
DEFINITION Oryza sativa (japonica cultivar-group) chromosome 10, section 28 of
77 of the complete sequence.
ACCESSION AE017074 AE016959
VERSION AE017074.1 GI:31431051
KEYWORDS Oryza sativa (japonica cultivar-group)
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartioideae; Oryzeae; Oryza.
1 (bases 1 to 304110)
REFERENCE
AUTHORS The Rice Chromosome 10 Sequencing Consortium
CONSTRM In-depth view of structure, activity, and evolution of rice
TITLE chromosome 10
JOURNAL Science 300, 1566-1569 (2003)
REFERENCE 2 (bases 1 to 304110)
AUTHORS Buell, C.R., Wing, R.A., McCombie, W.R., Messing, J. and Yuan, Q.
TITLE Direct Submission
JOURNAL Submitted (05-MAY-2003) The Institute for Genomic Research, 9712
Medical Center Dr., Rockville, MD 20850, USA
COMMENT This is the pseudomolecule for rice chromosome 10, which was
constructed by resolving discrepancies between overlapping BACs,
trimming the overlap regions, and linking the unique sequences to
form a contiguous sequence. Genes in individual BAC clone were
identified by a combination of several methods: Gene prediction
programs, searches of the complete sequence against a peptide
database and EST databases. Genes with similarity to other proteins
are named after the database hits. Genes without significant
peptide similarity but with EST similarity are named as unknown
proteins. Genes without protein or EST similarity, that are
predicted by more than two gene prediction programs over most of
their length are annotated as hypothetical proteins. Genes
encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy,
http://genome.wustl.edu/eddy/tRNAscan-SE/).
FEATURES
source
1. .304110
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/chromosome="10"
misc_feature
<1. .98612
/note="Chromosome Sequence Derivation: nucleotide sequence
in this region was derived from BAC clone OSJNBb0079E01
(GB:AC119149)."
gene
/locus_tag="OSJNBb0079E01.13"
/note="Similar to sp Q05654 RDPO_SCHPO Retrotransposable
element Tf2 155 kDa protein E value e 135; contains pfam
domain: rvt Reverse transcriptase (RNA dependent D e value
3 4e 48; contains pfam domain: rve Integrase core domain e
value 1 3e 34; contains pfam
complement (join(<12208. .14798,15340. .15431,15774. .>17629))
/locus_tag="OSJNBb0079E01.13"
mRNA
complement (join(12208. .14798,15340. .15431,15774. .17629))
CDS
/locus_tag="OSJNBb0079E01.13"

Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	20.93%	Indels:	0
DB:	8	Gaps:	0
US-10-624-218-3 (1-43) x AE017074 (1-304110)			
QY	25	ArgValGlnArgLeuArgGlnArgVal	33
DB	250467	AGGGTTCAAAGGTTAAGGAACAGATG	250441
RESULT 9			
AC105818/c			
LOCUS	AC105818	305902 bp	DNA linear HTG 11-OCT-2002
DEFINITION	Rattus norvegicus clone CH230-261KS, *** SEQUENCING IN PROGRESS		
ACCESSION	AC105818.4 GI:23603214		
VERSION	HTG; HTGS PHASE1; HTGS DRAFT; HTGS_ENRICHED.		
KEYWORDS	Rattus norvegicus (Norway rat)		
SOURCE	Rattus norvegicus		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.		
REFERENCE	1 (bases 1 to 305902)		
AUTHORS	Muzny,D., Marie,, Metzker,M., Lee,, Abramson,S., Adams,C., Alder,J., Allen,C., Allen,H., Alsbrooks,S., Amin,A., Angiano,D., Anyalebechi,V., Ayodeji,M., Baca,E., Baden,H., Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F., Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M., Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E., Cardenas,V., Carter,K., Cavazos,I., Cesar,H., Center,A., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J., Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L., Davila,M.B., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D., Delgado,O., Denson,S., Deramo,C., Ding,Y., Dirh,H., Divya,K., Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K., Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G., Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P., Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M., Gebregregis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W., Gunaratne,P., Haaland,W., Hamill,C., Hamilton,C., Hamilton,K., Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J., Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M., Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A., Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A., Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C., Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J., Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J., Lu,X., Ma,J., Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A., Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E., Mawhney,S., McLeod,M.P., McNeill,T.Z., Meenen,E., Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S., Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L., Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Nwackeleme,O., Okwuon,G., Olarpunsagoon,A., Pal,S., Parks,K., Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C., Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L., Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F., Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J., Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H., Shetty,J., Shvartsbeyn,A., Sison,I., Sitter,C.D., Smajs,D., Sneed,A., Sodergren,B., Song,X.-Z., Soralle,R., Sosa,J., Steidle,M., Strong,R., Sutton,A., Svatek,A., Taboz,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K., Valas,R., Vera,V., Villalana,D., Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F., Williams,G., Willson,R., Wleczek,R., Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,X., Zhou,X., Zhao,S., Zhao,S., Dunn,D., von Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,		

Weinstock,G. and Gibbs,R.A.
 Direct Submission
 Unpublished
 2 (bases 1 to 305902)
 Worley,K.C.
 Direct Submission
 Submitted (10-JAN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 305902)
 Rat Genome Sequencing Consortium.
 Direct Submission
 Submitted (11-OCT-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 On Oct 9, 2002 this sequence version replaced gi:23195953.
 The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: GNVB
 Center clone name: CH230-261KS
 ----- Summary Statistics
 Assembly program: Phrap; version 0.990329
 Consensus quality: 230544 bases at least Q40
 Consensus quality: 237984 bases at least Q30
 Consensus quality: 243263 bases at least Q20
 Estimated insert size: 235445; sum-of-contigs estimation
 Quality coverage: 5x in Q20 bases; sum-of-contigs estimation

 * NOTE: Estimated insert size may differ from sequence length
 (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently consists of 21 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will be preserved.

1	52223:	contig of 52223 bp in length
*	52224:	gap of unknown length
*	52323:	contig of 4257 bp in length
*	56580:	gap of unknown length
*	56581:	contig of 65011 bp in length
*	56681:	contig of 24971 bp in length
*	121692:	contig of 24971 bp in length
*	121792:	contig of 24971 bp in length
*	146763:	gap of unknown length
*	146862:	contig of 14834 bp in length
*	161696:	gap of unknown length
*	161797:	contig of 42128 bp in length
*	203925:	contig of 13457 bp in length
*	204025:	contig of 13457 bp in length
*	217482:	contig of 12061 bp in length
*	217582:	contig of 12061 bp in length
*	229742:	gap of unknown length
*	229743:	contig of 16203 bp in length
*	245946:	gap of unknown length
*	259797:	contig of 13752 bp in length

* 259798 259997: gap of unknown length
* 259898 275971: contig of 16074 bp in length
* 275972 276071: gap of unknown length
* 276072 277267: contig of 1196 bp in length
* 277268 277367: gap of unknown length
* 277368 278401: contig of 1034 bp in length
* 278402 278501: gap of unknown length
* 278502 280002: contig of 1501 bp in length
* 280003 281102: gap of unknown length
* 281103 281712: contig of 1610 bp in length
* 281713 283247: contig of unknown length
* 283248 283347: gap of unknown length
* 283348 284597: contig of 1250 bp in length
* 284598 284697: gap of unknown length
* 284698 286060: contig of 1363 bp in length
* 286061 286160: gap of unknown length
* 286161 287708: contig of 1548 bp in length
* 287709 287808: gap of unknown length
* 287809 289919: contig of 2111 bp in length
* 289920 290019: gap of unknown length
* 290020 305902: contig of 15883 bp in length.

FEATURES

source

1. 305902
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-261K5"

misc_feature

6728..7587
/note="clone boundary
clone_end:Sp6
site:MboI

misc_feature

end sequence:RXADI63TV"
complement (8531..9375)
/note="clone boundary
clone_end:T7
site:MboI

misc_feature

end sequence:RXADI63TV"
85409..86783
/note="wgs_end_extension
clone_end:T7"

misc_feature

135085..137212
/note="wgs_end_extension
clone_end:T7"

misc_feature

139392..141741
/note="wgs_end_extension
clone_end:T7"

misc_feature

146863..150200
/note="wgs_end_extension
clone_end:T7"

misc_feature

161797..163126
/note="wgs_end_extension
clone_end:T7"

misc_feature

201353..203924
/note="wgs_end_extension
clone_end:T7"

misc_feature

204025..205586
/note="wgs_end_extension
clone_end:T7"

misc_feature

229743..230901
/note="wgs_end_extension
clone_end:T7"

misc_feature

231403..232682
/note="wgs_end_extension
clone_end:T7"

misc_feature

274525..275971
/note="wgs_end_extension

Alignment Scores:

Pred. No.: 336 Length: 305902
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0

Query Match: 20.93% Indels: 0
DB: Gaps: 0
US-10-624-218-3 (1-43) x AC105818 (1-305902)
Qy 11 GluAsnThrAlaLeuGluThrArgVal 19
|||||
Db 115625 GAAACACTGCTCCCTGGAGACCCGAGTT 115599
|||||
RESULT 10
LOCUS G73580 440 bp DNA linear STS 28-DEC-2002
DEFINITION OMM1350 Rainbow Trout Blood Oncorhynchus mykiss STS genomic,
sequence tagged site.
ACCESSION G73580
VERSION G73580.1 GI:19111147
KEYWORDS STS.
SOURCE Oncorhynchus mykiss (rainbow trout)
ORGANISM Oncorhynchus mykiss
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei;
Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
REFERENCE 1 (bases 1 to 440)
AUTHORS Palti,Y., Fincham,M.R. and Rexroad,C.E. III.
TITLE Characterization of 38 polymorphic microsatellite markers for
rainbow trout (Oncorhynchus mykiss)
JOURNAL Mol. Ecol. Notes 2 (4), 449-452 (2002)
COMMENT

Contact: Yniv Palti
National Center for Cool and Cold Water Aquaculture
Agricultural Research Service, United States Department of
Agriculture
11876 Leetown Rd, Kearneysville, WV 25430, USA
Tel: 304-724-8340
Fax: 304-725-0351
Email: ypalti@cccwa.ars.usda.gov
Primer A: GGTGGAATGGGTGACATGAT
Primer B: CTCGTCCCTAACCTTTACCCCTG
STS size: 140
PCR Profile:

Presoak: 94 degrees C for 10.00 minutes
Denaturation: 94 degrees C for 30 seconds
Annealing: 58 degrees C for 30 seconds
Polymerization: 72 degrees C for 30 seconds
PCR Cycles: 35
Thermal Cycler: MJ Research PTC 200
Protocol:
Template: 30-100 pg
Primer: each 10 uM
dNTPs: each 200 uM
Taq Polymerase: 0.05 units/ul
Total Vol: 12 ul

Buffer:
Amplitaq Gold Buffer II w/o MgCl2
MgCl2: 2.5 mM.

FEATURES

source

Location/Qualifiers
1..440
/organism="Oncorhynchus mykiss"
/mol_type="genomic DNA"
/db_xref="taxon:8022"
/clone_lib="Rainbow Trout Blood"
/dev_stage="adult"
/note="Vector: pUC19; V-type: Plasmid; Genomic Hind III
fragments cloned into pUC19"

STS

primer_bind 226..365
primer_bind 226..246
complement(342..365)

ORIGIN

Alignment Scores:
Pred. No.: 19 Length: 440
Score: 8.00 Matches: 8

Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	18.60%	Indels:	0
DB:	11	Gaps:	0
US-10-624-218-3 (1-43) x G73580 (1-440)			
QY	5 AlaAlaPheLeuGluArgGluAsn 12		
Db	74 GCCGCTTCCTGGAGAGAGAGAAC 97		
RESULT 11			
WHTWALI5A/c			
LOCUS	WHTWALI5A	520 bp	mRNA linear
DEFINITION	Triticum aestivum protein of unknown function (wallis) mRNA, complete cds.		
ACCESSION	L11882		
VERSION	L11882.1	GI:170794	
KEYWORDS	Triticum aestivum (bread wheat)		
SOURCE	Triticum aestivum		
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae; Triticeae; Triticum.		
REFERENCE	1 (bases 1 to 520)		
AUTHORS	Snowden,K.C. and Gardner,R.C.		
TITLE	Five genes induced by aluminum in wheat (Triticum aestivum L.) roots		
JOURNAL	Plant Physiol. 103 (3), 855-861 (1993)		
MEDLINE	94294563		
PUBMED	8022939		
REFERENCE	2 (bases 1 to 520)		
AUTHORS	Snowden,K.C.		
TITLE	Direct Submission		
JOURNAL	Submitted (05-MAR-1993) Kimberley Snowden, Centre for Gene Technology, Dept of Cellular & Molecular Biology, University of Auckland, Private Bag, Auckland, New Zealand		
COMMENT	Original source text: Triticum aestivum (cultivar Warigal) cDNA to mRNA.		
FEATURES	Location/Qualifiers		
source	1..520		
	/organism="Triticum aestivum"		
	/mol_type="mRNA"		
	/cultivar="warigal"		
	/db_xref="taxon:4565"		
gene	tissue_type="aluminum-treated root tip"		
	1..520		
mRNA	/gene="wallis"		
	<1..520		
5'UTR	/gene="wallis"		
	<1..72		
CDS	/gene="wallis"		
	73..342		
	/gene="wallis"		
	/function="unknown"		
	/codon_start=1		
	/protein_id="AA050850.1"		
	/db_xref="GI:170795"		
	/translation="MKGTKLAAILLOAVLVMGLLSHVADFFPKCCNCRSFSQVDV		
	CDDAHPKCPQGSACRVVSTSPFWMRCDMKSTVDGTCGGPCKY"		
3'UTR	343..520		
	/gene="wallis"		
polyA_site	520		
	/gene="wallis"		
ORIGIN			
Alignment Scores:	21.9	Length:	520
Pred. No.:	8.00	Matches:	8
Score:	8.00	Conservative:	0
Percent Similarity:	100.00%	Mismatches:	0
Best Local Similarity:	100.00%	Indels:	0
Query Match:	18.60%	Gaps:	0
DB:	8		
US-10-624-218-3 (1-43) x WHTWALI5A (1-520)			
QY	7 PheLeuGluArgGluAsnThrAla 14		
Db	494 TTTCTCGAGAGAGAAAAACACACGCG 471		
RESULT 12			
AR505926			
LOCUS	AR505926	891 bp	DNA linear
DEFINITION	Sequence 10886 from patent US 6703491.		
ACCESSION	AR505926		
VERSION	AR505926.1	GI:52441401	
KEYWORDS	Unknown.		
SOURCE	Unknown.		
ORGANISM	Unclassified.		
REFERENCE	1 (bases 1 to 891)		
AUTHORS	Homburger,S.A., Ebens,A.J. Jr., Erickson,C.S., Francis-Lang,H.L., Margolis,J.S., Reddy,B.P., Ruddy,D.A. and Buchman,A.R.		
TITLE	Drosophila sequences		
JOURNAL	Patent: US 6703491-A 10886 09-MAR-2004;		
FEATURES	Location/Qualifiers		
source	1..891		
	/organism="unknown"		
	/mol_type="genomic DNA"		
ORIGIN			
Alignment Scores:	34.2	Length:	891
Pred. No.:	8.00	Matches:	8
Score:	8.00	Conservative:	0
Percent Similarity:	100.00%	Mismatches:	0
Best Local Similarity:	100.00%	Indels:	0
Query Match:	18.60%	Gaps:	0
DB:	6		
US-10-624-218-3 (1-43) x AR505926 (1-891)			
QY	15 LeuGluThrArgValAlaGluLeu 22		
Db	273 CTGAAACCCGTGTGGCGGAGCTT 296		
RESULT 13			
DRU96848			
LOCUS	DRU96848	1459 bp	mRNA linear
DEFINITION	Danio rerio thyrotroph embryonic factor beta (tefbeta) mRNA, complete cds.		
ACCESSION	U96848		
VERSION	U96848.1	GI:2828804	
KEYWORDS	Danio rerio (zebrafish)		
SOURCE	Danio rerio		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.		
REFERENCE	1 (bases 1 to 1459)		
AUTHORS	Xu,X., Liu,L., Wong,K.C. and Ge,R.		
TITLE	Cloning and characterization of two isoforms of the zebrafish thyrotroph embryonic factor (tef alpha and tefbeta)		
JOURNAL	Biochim. Biophys. Acta 1395 (1), 13-20 (1998)		
MEDLINE	98094234		
PUBMED	9434145		
REFERENCE	2 (bases 1 to 1459)		
AUTHORS	Xu,X., Liu,L., Wong,K.C.Y. and Ge,R.		
TITLE	Direct Submission		
JOURNAL	Submitted (08-APR-1997) School of Biological Science, National University of Singapore, Lower Kent Ridge Road, Singapore 119260, Republic of Singapore		
FEATURES	Location/Qualifiers		
source	1..1459		
	/organism="Danio rerio"		
	/mol_type="mRNA"		
	/db_xref="taxon:7955"		

gene 1. 1459
/gene="tefbeta"
72. .953
/gene="tefbeta"
/note="PAR subfamily"
/codon_start=1
/product="thyrotroph embryonic factor beta"
/protein_id="AAC00211.1"
/db_xref="GI:2828805"
/translation="MSSEIPFIFKALLEYPSLPISIDNENDKEKLPESVSGVSM
GPSALPAPAIWKPTPYDQDTHFLEYMDLREFLMENGIAAENQKSEKENIOLTAE
EPSTASAVKATPATVILPVALDPCREEVITITSSSSADNKSEENRMTDPINPDE
IEVDNPEPDDLVLSIPGGELFDRKRFSEELKQPMIKAKKVFVPEDQKDD
KYWRKKNVNAKRSRDLKENQITVRAAFLENSALRQEVABLKDQFGRCKNT
VARYEYKYGALGPEDV"

CDS

ORIGIN

Alignment Scores:
Pred. No.: 51.5 Length: 1459
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 18.60% Indels: 0
DB: 5 Gaps: 0

US-10-624-218-3 (1-43) x DRU96848 (1-1459)

Qy 5 AlaAlaPheLeuGluArgGluAsn 12
Db 816 CCGCGATTCTGGAGAGAGAAAC 839

RESULT 14

AY060258 1594 bp mRNA linear INV 05-NOV-2001

LOCUS Drosophila melanogaster GH08338 full length cDNA.

DEFINITION AY060258

ACCESSION AY060258.1 GI:16648063

VERSION

KEYWORDS FLI CDNA.

SOURCE Drosophila melanogaster (fruit fly)

ORGANISM Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.

1 (bases 1 to 1594)

Stapleton,M., Brokstein,P., Hong L., Agbayani,A., Carlson,J.,
Champe,M., Chavez,C., Dorsett,V., Farfan,D., Frise,E., George,R.,
Gonzalez,W., Guarin,H., Li,P., Liao,G., Miranda,A., Mungall,C.J.,
Nuncoo,J., Pacleb,J., Paragas,V., Park,S., Phouanavong,S., Wan,K.,
Yu,C., Lewis,S.E., Rubin,G.M. and Celnikier,S.

Direct Submission

Submitted (29-OCT-2001) Berkeley Drosophila Genome Project,
Lawrence Berkeley National Laboratory, One Cyclotron Road,
Berkeley, CA 94720, USA

Sequence submitted by:
Berkeley Drosophila Genome Project
Lawrence Berkeley National Laboratory
Berkeley, CA 94720

This clone was sequenced as part of a high-throughput process to
sequence clones from Drosophila Gene Collection 1 (Rubin et al.,
Science 2000). The sequence has been subjected to integrity checks
for sequence accuracy, presence of a polyA tail and contiguity
within 100 kb in the genome. Thus we believe the sequence to
reflect accurately this particular cDNA clone. However, there are
artifacts associated with the generation of cDNA clones that may
have not been detected in our initial analyses such as internal
priming, priming from contaminating genomic DNA, retained introns
due to reverse transcription of unspliced precursor RNAs, and
reverse transcriptase errors that result in single base changes.
For further information about this sequence, including its location
and relationship to other sequences, please visit our Web site
(http://fruitfly.berkeley.edu) or send email to
cdna@fruitfly.berkeley.edu.

Location/Qualifiers

FEATURES

source 1. 1594
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/strain="Y; cn bw sp"
/db_xref="taxon:7227"
/map="66A22-66A22"
1. 1594
/gene="RNaseX25"
/note="alignment with genomic scaffold AE003557"
/db_xref="FLYBASE:FBgn0010406"
199. .972
/gene="RNaseX25"
/note="Longest ORF"
/codon_start=1
/product="GH08338p"
/protein_id="AAL25297.1"
/db_xref="GI:16648064"
/db_xref="FLYBASE:FBgn0010406"
/translation="MSVQDNWDVLIETQWPVTTTCYHVRNPDQCSLPQKKEFWT
IHGWPTKLHQMGNFNCNSANFDPKLNPIEDRLTFWPDCLKGMDSTELWKHEWQK
HGTCAMLVEELDNELKYEQLTWREYIMSRILDASDIHDPDSNNTVAAINAIWKAL
GKNPSIRCLYDKHGHSYLSIRICFSKSLRIDCDGKQGDVAVPGVPGGTITTNCH
IGSLVHPSLVPPPLQKSHWKPLVNVYKLIQLMWFTL"

gene

CDS

ORIGIN

Alignment Scores:
Pred. No.: 55.4 Length: 1594
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 18.60% Indels: 0
DB: 3 Gaps: 0

US-10-624-218-3 (1-43) x AY060258 (1-1594)

Qy 14 AlaLeuGluThrArgValAlaGlu 21
Db 480 GCTTTGGAAACACGAGTGGCAGAA 503

RESULT 15

AX755546 1620 bp DNA linear PAT 23-JUN-2003

LOCUS Sequence 247 from Patent WO03000905.

DEFINITION AX755546

ACCESSION AX755546

VERSION AX755546.1 GI:32167856

KEYWORDS

SOURCE Oryza sativa

ORGANISM Oryza sativa

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.

1

REFERENCE Zhu,T., Cheng,W., Briggs,S., Cooper,B., Goff,S.A., Moughamer,T.,
Glazebrook,J., Katagiri,F., Kreps,J., Provart,N. and Ricke,D.

AUTHORS

TITLE Identification and characterization of plant genes

JOURNAL Patent: WO 03000905-A 247 03-JAN-2003;
Syngenta Participations AG (CH)

FEATURES Location/Qualifiers

source 1. 1620
/organism="Oryza sativa"
/mol_type="unassigned DNA"
/db_xref="taxon:4530"

ORIGIN

Alignment Scores:
Pred. No.: 56.2 Length: 1620
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 18.60% Indels: 0
DB: 6 Gaps: 0

US-10-624-218-3 (1-43) x AX755546 (1-1620)

Thu Jul 28 11:34:43 2005

Qy 25 ArgValGlnArgLeuArgAsnArg 32
Db 1252 AGAGTTCAGAGATTGAGANAATAGA 1275

Search completed: July 27, 2005, 13:56:15
Job time : 1909.92 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 27, 2005, 10:05:45 ; Search time 344.597 Seconds
(without alignments)
738.685 Million cell updates/sec

Title: US-10-624-218-3

Perfect score: 43

Sequence: 1 LEIEAFLERNTAETRA.....QVORLNRVSYQTRYGPL 43

Scoring table:

OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Word size: 1

Total number of hits satisfying chosen parameters: 8774676

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US10624218/runat_26072005_121435_3677/app.query.fasta_1.796
-DB=N Geneseq_16Dec04 -QWMT=fastap -SUFFIX=olip2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10624218 QCEN 1.1 1241 @runat_26072005_121435_3677 -NCPU=6 -ICPU=3
-NO_MMALP -LARGQUERY -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAEXT=60 -FGAPOP=6
-FGAEXT=7 -YGAPOP=60 -YGAEXT=60 -DELOP=6 -DELEXT=7

Database :

N Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	8	18.6	1273	ADJ39778	Adj39778 Plant cDN
2	8	18.6	1620	ADC07981	Adc07981 Rice DNA
3	8	18.6	1659	ABl13829	Abi13829 Drosophil
4	8	18.6	1796	ABl09365	Abi09365 Drosophil
5	8	18.6	2000	ADC08503	Adc08503 Rice DNA

6	8	18.6	2532	10	ADC30153	Adc30153 Human nov
7	8	18.6	2775	5	ABX71223	Abx71223 Human bra
8	8	18.6	2829	13	ADQ86932	Adq86932 Human tum
9	8	18.6	3489	4	AAS54094	Aas54094 Pseudomon
10	8	18.6	3489	8	ACA42240	Aca42240 Prokaryot
11	8	18.6	3796	4	ABL09364	Abi09364 Drosophil
12	8	18.6	3858	4	ABL13828	Abi13828 Drosophil
13	8	18.6	4149	12	ADI45378	Adi45378 Rice leop
14	8	18.6	5186	4	ABL13686	Abi13686 Drosophil
15	8	18.6	6155	4	ABL09166	Abi09166 Drosophil
16	7	16.3	40	2	AAZ23466	Aaz23466 Plant rd2
17	7	16.3	40	12	ADH75588	Adh75588 PCR prime
18	7	16.3	50	4	AAL34231	Aal34231 Human SNP
19	7	16.3	241	4	AAF17869	Aaf17869 Human bre
20	7	16.3	241	4	AAF47299	Aaf47299 Human bre
21	7	16.3	241	6	ABS63900	Abs63900 Human bre
22	7	16.3	241	10	ABT33112	Abt33112 Human tum
23	7	16.3	241	11	ADL93019	Adl93019 Human bre
24	7	16.3	241	12	ADE44309	Ad44309 Human CDN
25	7	16.3	254	6	ABL78740	Abi78740 Human ova
26	7	16.3	254	12	ADQ04143	Adq04143 Maize tra
27	7	16.3	281	6	ABL77738	Abi77738 Human ova
28	7	16.3	292	6	ABL77752	Abi77752 Human ova
29	7	16.3	307	12	ACH92882	Ach92882 Human gen
30	7	16.3	341	12	ADP92480	Adp92480 Cotton ex
31	7	16.3	342	5	AAH65976	Aah65976 C Glutami
32	7	16.3	357	5	AAH65974	Aah65974 C Glutami
33	7	16.3	357	8	ACA00290	Aca00290 C. glutam
34	7	16.3	360	9	ACH31454	Ach31454 Human bon
35	7	16.3	361	8	ABX55084	Abx55084 Bovine ES
36	7	16.3	450	8	ACC41446	Acc41446 Perennial
37	7	16.3	457	9	ACH28458	Ach28458 Human adu
38	7	16.3	465	6	ABL64140	Abi64140 Stomach c
39	7	16.3	472	3	AAC00941	Aac00941 Human sec
40	7	16.3	493	5	AAS05589	Aas05589 Mammalian
41	7	16.3	520	8	ACC41424	Acc41424 Perennial
42	7	16.3	524	12	ACH79182	Ach79182 Human gen
43	7	16.3	526	8	ACC41444	Acc41444 Perennial
44	7	16.3	527	8	ACC41445	Acc41445 Perennial
45	7	16.3	532	8	ACC41443	Acc41443 Perennial

ALIGNMENTS

RESULT 1
ADJ39778
ID ADJ39778 standard; cDNA; 1273 BP.

AC ADJ39778;

DT 06-MAY-2004 (first entry)

DE Plant cDNA #778.

XX Plant; gene; ss; transcription; plant genome augmentation; cereal;
KW soybean; alfalfa; sunflower; canola; cotton; peanut; tobacco; sugar beet;
KW maize; barley; sorghum; rice; wheat; crop plant; insecticide resistance;
KW stress tolerance; salt tolerance; cold tolerance; drought tolerance;
KW plant nutrition; apical dominance; dwarfism; early flowering; antiviral;
KW antifungal.

XX Eukaryota.

XX US2004016025-A1.

PD 22-JAN-2004.

XX 26-SEP-2002; 2002US-00260238.

XX 26-SEP-2001; 2001US-0325277P.

PR 26-SEP-2001; 2001US-0325448P.

XX 04-APR-2002; 2002US-0370620P.

PA (BUDW/) BUDWORTH P.
PA (MOUG/) MOUGHAMER T.
PA (BRIG/) BRIGGS S P.
PA (COOP/) COOPER B.
PA (GLAZ/) GLAZEBROOK J.
PA (GOFF/) GOFF S A.
PA (KATA/) KATAGIRI F.
PA (KREP/) KREPS J.
PA (PROV/) PROVART N.
PA (RICK/) RICHE D.
PA (ZHUT/) ZHU T.
XX
XX Budworth P, Moughamer T, Briggs SP, Cooper B, Glazebrook J;
PI Goff SA, Katagiri F, Kreps J, Provart N, Ricke D, Zhu T;
XX
XX WPI; 2004-190374/18.
XX
XX New rice promoter, useful for manipulating crop plants to alter or
PT improve phenotypic characteristics e.g. produce large quantities of oil
PT or high nutritional value.
XX
XX Claim 71; SEQ ID NO 778; 230pp; English.
XX
XX The invention relates to plant nucleotide sequences that direct seed-,
CC leaf- and/or stem-, panicle-, root- or pollen-specific or -preferential
CC or constitutive transcription of an operatively linked nucleic acid
CC segment. The invention also relates to a method for augmenting a plant
CC genome and a method of identifying a gene, where its expression is
CC altered in the seed, leaf, stem, panicle, pollen, root or is constitutive
CC in a plant cell. The plant is a cereal, e.g. soybean, alfalfa, sunflower,
CC canola, cotton, peanut, tobacco or sugar beet, preferably maize, barley,
CC sorghum, rice or wheat. The polynucleotides and the polypeptides they
CC encode are useful for manipulating crop plants to alter or improve
CC phenotypic characteristics, to produce large quantities of oil or
CC proteins, to incur resistance to insecticides, viruses or fungi, and to
CC incur stress tolerance (e.g. salt, cold or drought) to ensure the plants
CC have a high nutritional value with reduced apical dominance or dwarfism,
CC early flowering or altered metabolic pathways. This sequence represents a
CC plant nucleic acid of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification but was obtained in
CC electronic format directly from USPTO at seqdata.uspto.gov/sequence.html.
XX
XX Sequence 1273 BP; 330 A; 229 C; 334 G; 320 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 90.4 Length: 1273
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 18.60% Indels: 0
DB: 12 Gaps: 0

US-10-624-218-3 (1-43) x ADJ39778 (1-1273)

QY 25 ArgValGlnArgLeuArgAsnArg 32
DB 331 AGGGTTCAGAGATTGAGGACAGAGA 414

RESULT 2
ADCO7981
ID ADCO7981 standard; DNA; 1620 BP.
XX
XX ADCO7981;
XX
XX 18-DEC-2003 (first entry)
XX
XX Rice DNA sequence Seq ID247 related to grain filling.
XX
XX plant biotechnology; carbohydrate synthesis; carbohydrate metabolism;
KW carbohydrate degradation; carbohydrate; plant grain; grain filling; corn;
KW tomato; banana; canola; cotton; peanut; sorghum; tobacco; sugarbeet;
KW wheat; rice; protein; oil; starch; fibre; moisture content; cereal grain;

KW gene; ds; plant.
XX
XX Oryza sativa.
XX
XX WO2003000905-A2.
XX
XX 03-JAN-2003.
XX
XX 21-JUN-2002; 2002WO-IB002450.
XX
XX 22-JUN-2001; 2001US-0300112P.
XX 26-SEP-2001; 2001US-0325277P.
XX 20-DEC-2001; 2001US-0342327P.
XX
XX (SYGN) SYNGENTA PARTICIPATIONS AG.
XX
XX Zhu T, Cheng W, Briggs S, Cooper B, Goff SA, Moughamer T;
PI Glazebrook J, Katagiri F, Kreps J, Provart N, Ricke D;
XX
XX WPI; 2003-229341/22.
XX P-PSDB; ADCO7982.
XX
XX New plant genes encoding polypeptides having an activity involved in or
PT associated with the synthesis, metabolism or degradation of carbohydrates
PT in the plant grain useful in generating plants having improved
PT nutritional properties.
XX
XX Claim 16; SEQ ID NO 247; 130pp; English.
XX
XX This invention, in the area of plant biotechnology, relates to novel
CC polynucleotides comprising a nucleotide sequence encoding a protein which
CC is involved in or associated with the synthesis, metabolism or
CC degradation of carbohydrates in the plant grain and the expression of
CC which is up-regulated during grain filling. The plant is selected from
CC corn, tomato, banana, canola, cotton, peanut, sorghum, tobacco,
CC sugarcane, wheat, and rice. The invention may be useful for the
CC improvement of protein, oil, starch, fibre and moisture content of the
CC cereal grains. In addition, carbohydrate levels may be modified to a more
CC desirable level using the present invention. The present sequence is a
CC DNA sequence encoding a rice protein of the invention. Note: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/publishedpat_sequences.
XX
XX Sequence 1620 BP; 550 A; 272 C; 368 G; 424 T; 0 U; 6 Other;

Alignment Scores:
Pred. No.: 113 Length: 1620
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 18.60% Indels: 0
DB: 10 Gaps: 0

US-10-624-218-3 (1-43) x ADCO7981 (1-1620)

QY 25 ArgValGlnArgLeuArgAsnArg 32
DB 1252 AGAGTTCAGAGATTGAGAAATAGA 1275

RESULT 3
ABLI3829
ID ABLI3829 standard; cDNA; 1659 BP.
XX
XX ABLI3829;
XX
XX 26-MAR-2002 (first entry)
XX
XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 35969.
XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ss.
KW

```
OS Drosophila melanogaster.
XX WO200171042-A2.
XX
XX 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US009231.
XX
XX 23-MAR-2000; 2000US-0191637P.
XX
XX 11-JUL-2000; 2000US-00614150.
XX
XX (PEKE ) PE CORP NY.
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
XX
XX WPI; 2001-656860/75.
XX
XX P-PSDB; ABB69726.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions.
XX
XX Claim 1; SEQ ID NO 35969; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-
XX ABB72072). The sequence data for this patent did not form part of the
XX printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX SQ Sequence 1659 BP; 467 A; 369 C; 347 G; 476 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 115 Length: 1659
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 18.60% Indels: 0
DB: 4 Gaps: 0
XX
US-10-624-218-3 (1-43) x ABL13829 (1-1659)
QY 14 AlaLeuGluThrArgValaGlu 21
Db 562 GCTTTGGAACACGAGTGCAGAA 585
XX
RESULT 4
ABL09365
ID ABL09365 standard; cDNA; 1796 BP.
XX
XX ABL09365;
XX
XX 26-MAR-2002 (first entry)
XX
XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 22577.
XX
XX Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical; gene; ss.
XX
XX Drosophila melanogaster.
XX
XX WO200171042-A2.
XX
XX 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US009231.
XX
XX 23-MAR-2000; 2000US-0191637P.
XX
XX 23-MAR-2000; 2000US-0191637P.
XX
XX (SYGN ) SYNGENTA PARTICIPATIONS AG.
XX
XX Zhu T, Cheng W, Briggs S, Cooper B, Goff SA, Moughamer T;
XX
PI
```

```
PR 11-JUL-2000; 2000US-00614150.
XX
XX (PEKE ) PE CORP NY.
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
XX
XX WPI; 2001-656860/75.
XX
XX P-PSDB; ABB65262.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions.
XX
XX Claim 1; SEQ ID NO 22577; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-
XX ABB72072). The sequence data for this patent did not form part of the
XX printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX SQ Sequence 1796 BP; 495 A; 465 C; 454 G; 382 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 124 Length: 1796
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 18.60% Indels: 0
DB: 4 Gaps: 0
XX
US-10-624-218-3 (1-43) x ABL09365 (1-1796)
QY 15 LeuGluThrArgValaGluLeu 22
Db 1178 CTGGAACCCGCTGTGGCGAGCTT 1201
XX
RESULT 5
ADC08503/c
ID ADC08503 standard; DNA; 2000 BP.
XX
XX ADC08503;
XX
XX 18-DEC-2003 (first entry)
XX
XX Rice DNA sequence Seq ID808 related to grain filling.
XX
XX plant biotechnology; carbohydrate synthesis; carbohydrate metabolism;
XX carbohydrate degradation; carbohydrate; plant grain; grain filling; corn;
XX tomato; banana; canola; cotton; peanut; sorghum; tobacco; sugarbeet;
XX wheat; rice; protein; oil; starch; fibre; moisture content; cereal grain;
XX gene; ds; plant.
XX
XX Oryza sativa.
XX
XX WO2003000905-A2.
XX
XX 03-JAN-2003.
XX
XX 21-JUN-2002; 2002WO-IB002450.
XX
XX 22-JUN-2001; 2001US-0300112P.
XX
XX 26-SEP-2001; 2001US-0325277P.
XX
XX 20-DEC-2001; 2001US-0342327P.
XX
XX (SYGN ) SYNGENTA PARTICIPATIONS AG.
XX
XX Zhu T, Cheng W, Briggs S, Cooper B, Goff SA, Moughamer T;
XX
PI
```


KW Human; gene; gene therapy; vaccine; disease treatment; detection; ss.
 XX Homo sapiens.
 OS WO200112659-A2.
 XX 22-FEB-2001.
 XX 18-AUG-2000; 2000WO-IB001496.
 XX 18-AUG-1999; 99US-0149499P.
 PR 28-SEP-1999; 99US-0156503P.
 XX (GEHU-) GERMAN HUMAN GENOME PROJECT.
 XX Wiemann S;
 XX WPI; 2001-327840/34.
 DR P-PSDB; ABUS2641.
 XX Nucleic acids having the sequences of clones isolated from libraries of
 PT different human tissues, useful in recombinant DNA methodologies.
 XX Claim 1; Page 162-163; 1095pp; English.
 PS This invention describes novel polynucleotides and polypeptides isolated
 CC from human cDNA libraries which can be used for gene therapy or in
 CC vaccines. The polynucleotides of the invention and antibodies encoded by
 CC them may be used in the prevention, diagnosis and treatment of diseases
 CC associated with inappropriate polypeptide expression. The products of the
 CC invention may also be used to identify modulators of expression and
 CC activity and to down regulate expression and activity. The antibodies of
 CC the invention may also be used as diagnostic agents for detecting the
 CC presence of polypeptides in samples. This sequence encodes a polypeptide
 CC described in the disclosure of the invention
 XX SQ Sequence 2775 BP; 622 A; 771 C; 931 G; 451 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 185 Length: 2775
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 18.60% Indels: 0
 DB: 5 Gaps: 0
 US-10-624-218-3 (1-43) x ABX71223 (1-2775)
 QY 23 ArgGlnArgValGlnArgLeuArg 30
 DB 1087 CGACAGCGAGTTTCAGCGACTTCGA 1110
 RESULT 8
 ID ADQ86932
 XX ADQ86932 standard; cDNA; 2829 BP.
 AC ADQ86932;
 XX 07-OCT-2004 (first entry)
 DT Human tumour-associated antigenic target (TAT) cDNA sequence #3807.
 DE human; tumour-associated antigenic target; TAT; cytostatic; gene therapy;
 KW cancer; cell proliferative disorder; gene; ss.
 XX Homo sapiens.
 OS WO2004060270-A2.
 XX 22-JUL-2004.
 PD 15-OCT-2003; 2003WO-US029126.
 XX

PR 18-OCT-2002; 2002US-0418988P.
 XX (GETH) GENENTECH INC.
 PA (WUTD/) WU T D.
 PA (ZHOU/) ZHOU Y.
 XX Wu TD, Zhou Y;
 XX WPI; 2004-534300/51.
 DR New nucleic acid molecule and encoded polypeptide, for diagnosing,
 XX preventing or treating cell proliferative disorders such as cancer.
 PS Claim 1; SEQ ID NO 3807; 5504pp; English.
 XX The present invention describes an isolated tumour-associated antigenic
 CC target (TAT) nucleic acid comprising: (a) any of 4622 nucleotide
 CC sequences (see SEQ ID NO:1 to 4622); (b) the full-length coding region of
 CC (a); (c) the complement of (a) or (b); (d) a sequence that hybridises to (a)-
 CC sequence identity to (a)-(c); or (e) a sequence that hybridises to (a)-
 CC nucleic acid; (2) a host cell comprising the above expression vector; (3)
 CC a process for producing a polypeptide; (4) an isolated polypeptide
 CC comprising: (a) an amino acid sequence encoded by any of the above
 CC nucleotide sequences; (b) an amino acid sequence encoded by the full-
 CC length coding region of the above nucleotide sequences; or (c) a sequence
 CC having at least 80% identical to (a) or (b); (5) a chimeric polypeptide
 CC comprising the above polypeptide fused to a heterologous polypeptide; (6)
 CC an isolated antibody that binds to the above polypeptide; (7) a process
 CC for producing the antibody; (8) an isolated oligopeptide that binds to
 CC the above polypeptide; (9) a tumour-associated antigenic target (TAT)
 CC binding organic molecule that binds to the above polypeptide; (10) a
 CC composition of matter comprising the above (chimeric) polypeptide,
 CC antibody, oligopeptide or TAT binding organic molecule, in combination
 CC with a carrier; (11) an article of manufacture comprising a container and
 CC the composition of matter contained within the container; (12) methods of
 CC inhibiting the growth of a cell that expresses the above protein, where
 CC the growth of the cell is at least in part dependent upon a growth
 CC potentiating effect of the above protein; (13) a method of
 CC therapeutically treating a mammal having a cancerous tumour comprising
 CC cells that express the above protein; (14) a method of determining the
 CC presence of a protein in a sample suspected of containing the protein
 CC described above; (15) methods of diagnosing the presence of a tumour in a
 CC mammal; (16) a method for treating or preventing a cell proliferative
 CC disorder associated with increased expression or activity of the above
 CC protein; and (17) a method of binding an antibody, oligopeptide or
 CC organic molecule to a cell that expresses the protein described above.
 CC The TAT sequences have cytostatic activities, and can be used in gene
 CC therapy. The composition and methods are useful for diagnosing,
 CC preventing or treating cancer. The composition is also used for preparing
 CC a medicament for the therapeutic treatment or diagnostic detection of a
 CC cell proliferative disorder or cancer. The present sequence represents a
 CC human TAT cDNA sequence from the present invention.
 XX SQ Sequence 2829 BP; 620 A; 784 C; 960 G; 465 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 188 Length: 2829
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 18.60% Indels: 0
 DB: 13 Gaps: 0
 US-10-624-218-3 (1-43) x ADQ86932 (1-2829)
 QY 23 ArgGlnArgValGlnArgLeuArg 30
 DB 1161 CGACAGCGAGTTTCAGCGACTTCGA 1184
 RESULT 9
 ID AAS54094
 XX AAS54094 standard; DNA; 3489 BP.

XX AAS54094;
AC
XX
DT 13-FEB-2002 (first entry)
XX
XX Pseudomonas aeruginosa DNA for cellular proliferation protein #225.
DE
XX Antisense; ds; prokaryotic cellular proliferation gene; antibiotic;
KW antibacterial; drug design.
XX
XX Pseudomonas aeruginosa.
OS
XX WO200170955-A2.
PN
XX 27-SEP-2001.
PD
XX
XX 21-MAR-2001; 2001WO-US009180.
PF
XX 21-MAR-2000; 2000US-0191078P.
PR 23-MAY-2000; 2000US-0206848P.
PR 26-MAY-2000; 2000US-0207727P.
PR 23-OCT-2000; 2000US-0242578P.
PR 27-NOV-2000; 2000US-0253625P.
PR 22-DEC-2000; 2000US-0257931P.
PR 16-FEB-2001; 2001US-0269308P.
XX
XX (ELIT-) ELITRA PHARM INC.
FA
XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
XX Yamamoto RT, Xu HH;
PI
XX WPI; 2001-611495/70.
XX P-PSDB; AAU36235.
DR
XX
XX New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids.
PT
XX Claim 27; SEQ ID NO 7731; 511pp; English.
PS
XX The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the genes,
CC their use in the discovery of novel antibiotics, the essential genes
CC themselves and the encoded proteins. The prokaryotes used are Escherichia
CC coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
CC Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
CC useful for the identification of potential new targets for antibiotic
CC development. The antisense nucleic acids can also be used to identify
CC proteins used in proliferation, to express these proteins, and to obtain
CC antibodies capable of binding to the expressed proteins. The proteins can
CC be used to screen compounds in rational drug discovery programmes. The
CC antisense nucleic acid sequence is also useful to screen for homologous
CC nucleic acids which are required for cell proliferation in a wide variety
CC of organisms. The present sequence encodes an essential prokaryotic
CC cellular proliferation protein. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 3489 BP; 697 A; 1090 C; 1228 G; 474 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 228 Length: 3489
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 18.60% Indels: 0
DB: 4 Gaps: 0

US-10-624-218-3 (1-43) x AAS54094 (1-3489)

QY 14 AlaLeuGluThrArgValAlaGlu 21
Db 1993 GCCCTGGAACCCGTCGCCGAA 2016

RESULT 10
ACA42240
ID ACA42240 standard; DNA; 3489 BP.
XX
AC ACA42240;
XX
DT 19-JUN-2003 (first entry)
XX
XX Prokaryotic essential gene #23897.
DE
XX Antisense; ds; prokaryotic essential gene; cell proliferation;
KW drug design; gene.
XX
OS Pseudomonas aeruginosa.
PN WO200277183-A2.
XX
XX 03-OCT-2002.
PD
XX 21-MAR-2002; 2002WO-US009107.
PF
XX 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
XX (ELIT-) ELITRA PHARM INC.
PA
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
XX Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
PI
XX WPI; 2003-029926/02.
XX P-PSDB; ABU38370.
DR
XX
XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
XX Claim 14; SEQ ID NO 30110; 1766pp; English.
PS
XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than S. aureus, S. typhimurium,
CC K. pneumoniae or P. aeruginosa. The present sequence is one of the target
CC prokaryotic essential genes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

Pred. No.:	246	Length:	3796
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	MinMatches:	0
Query Match:	18.60%	Indels:	0

ADI45378
ID ADI45378 standard; cDNA; 4149 BP.
XX
AC ADI45378;
XX
DT 22-APR-2004 (first entry)
XX
DE Rice isoprenoid biosynthesis-associated cDNA #155.
XX
KW Rice; isoprenoid biosynthesis; ss; gene; plant; isopentenyl diphosphate;
KW IPP; dimethylallyl alcohol; DMAPP; short-chain plastid prenyltransferase;
KW gibberellin; carotenoid; abscisic acid; tocopherol; plastoquinone;
KW phytylquinone; mevalonate pathway; phytosterol; brassinosteroid;
KW haeme; yield.
XX
OS Oryza sativa.
XX
PN US2004010815-A1.
XX
PD 15-JAN-2004.
XX
PF 26-SEP-2002; 2002US-00259194.
XX
PR 26-SEP-2001; 2001US-0325277P.
PR 04-APR-2002; 2002US-0370620P.
PR 04-APR-2002; 2002US-0370743P.
XX
PA (LANG/) LANGE B M.
PA (GHAS/) GHASSEMIA M.
PA (BRIG/) BRIGGS S P.
PA (COOP/) COOPER B.
PA (GLAZ/) GLAZEBROOK J.
PA (GOFF/) GOFF S A.
PA (KATA/) KATAGIRI F.
PA (KREP/) KREPS J.
PA (MOUG/) MOUGHAMER T.
PA (PROV/) PROVART N.
PA (RICK/) RICHKE D.
PA (ZHUT/) ZHU T.
XX
XX Lange BM, Ghassemian M, Briggs SP, Cooper B, Glazebrook J;
PI Goff SA, Katagiri F, Kreps J, Moughamer T, Provart N, Ricke D;
PI Zhu T;
XX
XX WPI; 2004-090562/09.
DR P-PSDB; ADI45379.
XX
XX New isolated polynucleotides and polypeptides associated with isoprenoid
PT synthesis in plants, useful for producing transgenic plants, for targeted
PT gene disruption, as well as markers or probes.
XX
XX Claim 1; SEQ ID NO 309; 117pp; English.
XX
XX The invention relates to a polynucleotide (or its complement, protein
CC encoding fragment or reverse complement), comprising a nucleotide
CC sequence encoding a polypeptide comprising an amino acid sequence
CC involved in or associated with the biosynthesis of isoprenoids in a rice
CC plant. Also included are an isolated polypeptide involved in or
CC associated with the biosynthesis of isoprenoids in a plant, an expression
CC cassette comprising the polynucleotide, a host cell comprising the
CC expression cassette, and a transgenic plant comprising the expression
CC cassette. The polypeptides and polynucleotides include those associated
CC with the biosynthesis of isopentenyl diphosphate (IPP) and dimethylallyl
CC alcohol (DMAPP), the biosynthesis of short-chain plastid
CC prenyltransferases, the biosynthesis of gibberellins, the biosynthesis of
CC carotenoids and/or abscisic acids, the biosynthesis of tocopherols,
CC plastoquinone and/or phytylquinone biosynthesis, the mevalonate pathway,
CC phytosterol and brassinosteroid metabolism, biosynthesis of ubiquinone,
CC biosynthesis of monoterpenes and sesquiterpenes, protein prenylation, and
CC biosynthesis of chlorophyll or haeme. Also disclosed are banana, wheat
CC and corn homologues of some of the rice polynucleotides. The
CC polynucleotides are useful for producing transgenic plants, where the

CC genome is augmented by a nucleic acid molecule of the invention, or in
CC which the corresponding gene has been disrupted, e.g. to result in a
CC loss, a decrease or an alteration in the function of the product encoded
CC by the gene. The plants may also have increased yields and/or produce a
CC better quality product than the corresponding wild-type plant. The
CC nucleic acid molecules are useful for targeted gene disruption, as well
CC as markers and probes. Note: The sequence data for this patent did not
CC form part of the printed specification, but was obtained in electronic
CC format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20040010815. The present sequence
CC is a Rice isoprenoid biosynthesis-associated cDNA of the invention.
XX
SQ Sequence 4149 BP; 1270 A; 774 C; 1049 G; 1056 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 267 Length: 4149
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 18.60% Indels: 0
DB: 12 Gaps: 0
US-10-624-218-3 (1-43) x ADI45378 (1-4149)
QY 25 ArgValGlnArgLeuArgAsnArg 32
DB 2146 AGCGTTCAAAGGTTGAGGAATAGA 2169
RESULT 14
ABLI13686/c
ID ABLI13686 standard; cDNA; 5186 BP.
XX
AC ABLI13686;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 35540.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ss.
XX
OS Drosophila melanogaster.
XX WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US009231.
XX
PR 23-MAR-2000; 2000US-0191637P.
PR 11-JUL-2000; 2000US-00614150.
XX
PA (PEKE) PE CORP NY.
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
PI WPI; 2001-656860/75.
DR P-PSDB; ABB69583.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX
PS Claim 1; SEQ ID NO 35540; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABLI16176-ABLI30511), expressed DNA
CC sequences (ABLI01840-ABLI16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the

CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 5186 BP; 1497 A; 1164 C; 1198 G; 1327 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 327 Length: 5186
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 18.60% Indels: 0
DB: 4 Gaps: 0

US-10-624-218-3 (1-43) x ABL13686 (1-5186)

Qy 14 AlaLeuGluThrArgValAlaGlu 21
|||||
Db 4630 GCTTTGGAAACGAGTGGCAGAA 4607

RESULT 15

ABL09166/c

ID ABL09166 standard; cDNA; 6155 BP.

XX ABL09166;

AC ABL09166;

XX 26-MAR-2002 (first entry)

XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 21980.

XX Drosophila; developmental biology; cell signalling; insecticide;

KW pharmaceutical; gene; ss.

XX Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US009231.

XX 23-MAR-2000; 2000US-0191637P.

XX 11-JUL-2000; 2000US-00614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

DR P-PSDB; ABB65063.

XX New isolated nucleic acid detection reagent for detecting 1000 or more

PT genes from Drosophila and for elucidating cell signalling and cell-cell

PT interactions.

XX Claim 1; SEQ ID NO 21980; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent

CC capable of detecting 1000 or more genes from Drosophila. The invention is

CC useful in developmental biology and in elucidating cell signalling and

CC cell-cell interactions in higher eukaryotes for the development of

CC insecticides, therapeutics and pharmaceutical drugs. The invention

CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA

CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-

CC ABB72072). The sequence data for this patent did not form part of the

CC printed specification, but was obtained in electronic format directly

CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX

SQ Sequence 6155 BP; 1715 A; 1424 C; 1493 G; 1523 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 383 Length: 6155

Score: 8.00 Matches: 8

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0
Query Match: 18.60% Indels: 0
DB: 4 Gaps: 0

US-10-624-218-3 (1-43) x ABL09166 (1-6155)

Qy 15 LeuGluThrArgValAlaGluLeu 22
|||||
Db 6051 CTGGAAACCCGTGTGGCGGAGCTT 6028

Search completed: July 27, 2005, 12:14:39
Job time : 353.597 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 27, 2005, 11:44:30 ; Search time 100.333 Seconds
(without alignments)
701.261 Million cell updates/sec

Title: US-10-624-218-3

Perfect score: 43

Sequence: 1 LEIBAEFLERNTALETRVA.....QRVQLRNRSQVTRYGPL 43

Scoring table:

OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Word size: 1

Total number of hits satisfying chosen parameters: 2398817

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US10624218/runat_26072005_121436_3711/app_query.fasta_1.796
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=olip2n.rn1 -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10624218@cgn_1_1_213@runat_26072005_121436_3711 -NCPU=6 -ICPU=3
-NO WMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : Issued Patents NA.*

1: /cgn2_6/ptodata/1/ina/5A COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PTCUT COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	8	18.6	891	4	US-09-270-767-10886
2	7	16.3	40	3	Sequence 10886, A
3	7	16.3	40	3	Sequence 28, Appl
4	7	16.3	241	4	US-09-601-802D-28
5	7	16.3	241	4	US-09-389-681-357
6	7	16.3	241	4	US-09-620-405B-357
7	7	16.3	241	4	US-09-433-826B-357
8	7	16.3	241	4	US-09-604-287A-357
9	7	16.3	241	4	US-09-834-759-357
10	7	16.3	241	4	US-09-590-751A-357
11	7	16.3	241	4	US-09-551-621-357
12	7	16.3	472	4	US-09-513-999C-939
			493	4	US-09-669-751-189

13	7	16.3	541	4	US-09-621-976-3315	Sequence 3315, Ap
14	7	16.3	548	4	US-09-513-999C-1569	Sequence 1569, Ap
15	7	16.3	573	4	US-09-621-976-3034	Sequence 3034, Ap
16	7	16.3	579	4	US-09-489-039A-4273	Sequence 4273, Ap
17	7	16.3	601	4	US-09-949-016-63485	Sequence 63485, A
18	7	16.3	601	4	US-09-949-016-133477	Sequence 133477, A
19	7	16.3	601	4	US-09-949-016-133478	Sequence 133478, A
20	7	16.3	601	4	US-09-949-016-133479	Sequence 133479, A
21	7	16.3	601	4	US-09-949-016-133480	Sequence 133480, A
22	7	16.3	601	4	US-09-949-016-152776	Sequence 152776, A
23	7	16.3	601	4	US-09-949-016-166493	Sequence 166493, A
24	7	16.3	601	4	US-09-949-016-166494	Sequence 166494, A
25	7	16.3	601	4	US-09-949-016-202640	Sequence 202640, A
26	7	16.3	654	4	US-09-252-991A-15023	Sequence 15023, A
27	7	16.3	666	4	US-09-252-991A-15023	Sequence 15023, A
28	7	16.3	702	4	US-09-489-039A-6546	Sequence 6546, Ap
29	7	16.3	784	4	US-09-205-258-101	Sequence 101, App
30	7	16.3	785	4	US-09-270-767-5707	Sequence 5707, Ap
31	7	16.3	785	4	US-09-270-767-20989	Sequence 20989, A
32	7	16.3	971	4	US-09-270-767-10525	Sequence 10525, A
33	7	16.3	1170	4	US-09-252-991A-14730	Sequence 14730, Ap
34	7	16.3	1236	4	US-09-252-991A-14730	Sequence 14730, A
35	7	16.3	1278	4	US-09-252-991A-14612	Sequence 14612, A
36	7	16.3	1323	4	US-09-489-039A-3491	Sequence 3491, Ap
37	7	16.3	1434	4	US-09-252-991A-3945	Sequence 3945, Ap
38	7	16.3	1458	4	US-09-252-991A-8297	Sequence 8297, Ap
39	7	16.3	1521	4	US-09-252-991A-3958	Sequence 3958, Ap
40	7	16.3	1599	1	US-08-285-440-11	Sequence 11, Appl
41	7	16.3	1599	1	US-08-630-349-11	Sequence 11, Appl
42	7	16.3	1677	1	US-08-285-440-12	Sequence 12, Appl
43	7	16.3	1677	1	US-08-630-349-12	Sequence 12, Appl
44	7	16.3	1785	4	US-09-489-039A-3451	Sequence 3451, Ap
45	7	16.3	1860	4	US-09-252-991A-5014	Sequence 5014, Ap

ALIGNMENTS

RESULT 1
US-09-270-767-10886
; Sequence 10886, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 10886
; LENGTH: 891
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-10886

Alignment Scores:
Pred. No.: 13.4 Length: 891
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 18.60% Indels: 0
DB: 4 Gaps: 0

US-10-624-218-3 (1-43) x US-09-270-767-10886 (1-891)

Qy 15 LeuGlutThrArgValaAlaGluLeu 22

Db 273 CTGGAACCCGTCGTGCGGAGCTT 296

RESULT 2

US-09-198-119C-28/c
; Sequence 28, Application US/09198119C
; Patent No. 6417428

```

; GENERAL INFORMATION:
; APPLICANT: Thomasow, Michael
; APPLICANT: Stockinger, Eric
; APPLICANT: Jaglo-Ottosen, Kirsten
; APPLICANT: Gilmour, Sarah
; APPLICANT: Zarka, Daniel
; APPLICANT: Jiang, Cai-Zhong
; TITLE OF INVENTION: Plant Having Altered Environmental Stress Tolerance
; FILE REFERENCE: 1917, 713 Seq List
; CURRENT APPLICATION NUMBER: US/09/198,119C
; PRIOR FILING DATE: 1998-11-23
; PRIOR APPLICATION NUMBER: US 08/706,270
; PRIOR FILING DATE: 1998-09-04
; PRIOR APPLICATION NUMBER: US 09/018,233
; PRIOR FILING DATE: 1998-02-03
; PRIOR APPLICATION NUMBER: US 09/017,816
; PRIOR FILING DATE: 1998-02-03
; PRIOR APPLICATION NUMBER: US 09/018,235
; PRIOR FILING DATE: 1998-02-03
; PRIOR APPLICATION NUMBER: US 09/017,575
; PRIOR FILING DATE: 1998-02-03
; PRIOR APPLICATION NUMBER: US 09/018,227
; PRIOR FILING DATE: 1998-02-03
; PRIOR APPLICATION NUMBER: US 09/018,234
; PRIOR FILING DATE: 1998-02-03
; NUMBER OF SEQ ID NOS: 95
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 28
; LENGTH: 40
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: PCR primer
US-09-198-119C-28

Alignment Scores:
Pred. No.:      8.66      Length:      40
Score:          7.00      Matches:      7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match:    16.28% Indels: 0
DB:             3       Gaps: 0

US-10-624-218-3 (1-43) x US-09-198-119C-28 (1-40)
Qy   10 ArgGlusnThrAlaLeuglu 16
     |||||
Db   30 AGAGAAACACAGCTTTGGAA 10

RESULT 3
US-09-601-802D-28/c
; Sequence 28, Application US/09601802D
; Patent No. 6706866
; GENERAL INFORMATION:
; APPLICANT: Thomasow, Michael
; APPLICANT: Stockinger, Eric
; APPLICANT: Jaglo-Ottosen, Kirsten
; APPLICANT: Gilmour, Sarah
; APPLICANT: Zarka, Daniel
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Zhang, James
; APPLICANT: Haake, Volker
; TITLE OF INVENTION: PLANT HAVING ALTERED ENVIRONMENTAL STRESS TOLERANCE
; FILE REFERENCE: 5144200201/MBI0029
; CURRENT APPLICATION NUMBER: US/09/601,802D
; CURRENT FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: 09/018,233
; PRIOR FILING DATE: 1998-02-03
; PRIOR APPLICATION NUMBER: 09/017,816
; PRIOR FILING DATE: 1998-02-03
; PRIOR APPLICATION NUMBER: 09/018,235
; PRIOR FILING DATE: 1998-02-03
; GENERAL INFORMATION:
; APPLICANT: Yuqiu, Jiang
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiaochun
; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND DIAGNOSIS OF BREAST CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.470C3
; CURRENT APPLICATION NUMBER: US/09/389,681A
; CURRENT FILING DATE: 1999-09-02
; NUMBER OF SEQ ID NOS: 463
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 357
; LENGTH: 241
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-389-681-357

Alignment Scores:
Pred. No.:      45.8      Length:      241
Score:          7.00      Matches:      7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match:    16.28% Indels: 0
DB:             4       Gaps: 0

US-10-624-218-3 (1-43) x US-09-389-681-357 (1-241)
Qy   15 LeuGluThrArgValalaGlu 21
     |||||
Db   208 CTGGAGACTCGGTGGCCGAG 188

RESULT 5
US-09-620-405B-357/c
; Sequence 357, Application US/09620405B
; Patent No. 6528054
; GENERAL INFORMATION:

```

; APPLICANT: Jiang, Yuqiu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.470C8
; CURRENT APPLICATION NUMBER: US/09/620,405B
; CURRENT FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 495
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 357
; LENGTH: 241
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-620-405B-357

Alignment Scores:
Pred. No.: 45.8 Length: 241
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 16.28% Indels: 0
DB: 4 Gaps: 0

US-10-624-218-3 (1-43) x US-09-620-405B-357 (1-241)

Qy 15 LeuGlutThrArgValAlaGlu 21
Db 208 CTGGAGACTCGGGTGGCCGAG 188

RESULT 6

US-09-433-826B-357/c
; Sequence 357, Application US/09433826B
; Patent No. 6579973
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
; FILE REFERENCE: 210121.470C4
; CURRENT APPLICATION NUMBER: US/09/433,826B
; CURRENT FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 474
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 357
; LENGTH: 241
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-433-826B-357

Alignment Scores:
Pred. No.: 45.8 Length: 241
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 16.28% Indels: 0
DB: 4 Gaps: 0

US-10-624-218-3 (1-43) x US-09-433-826B-357 (1-241)

Qy 15 LeuGlutThrArgValAlaGlu 21
Db 208 CTGGAGACTCGGGTGGCCGAG 188

RESULT 7

US-09-604-287A-357/c
; Sequence 357, Application US/09604287A

; Patent No. 6586572
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.470C7
; CURRENT APPLICATION NUMBER: US/09/604,287A
; CURRENT FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 489
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 357
; LENGTH: 241
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-604-287A-357

Alignment Scores:
Pred. No.: 45.8 Length: 241
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 16.28% Indels: 0
DB: 4 Gaps: 0

US-10-624-218-3 (1-43) x US-09-604-287A-357 (1-241)

Qy 15 LeuGlutThrArgValAlaGlu 21
Db 208 CTGGAGACTCGGGTGGCCGAG 188

RESULT 8

US-09-834-759-357/c
; Sequence 357, Application US/09834759
; Patent No. 6680197
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.470C9
; CURRENT APPLICATION NUMBER: US/09/834,759
; CURRENT FILING DATE: 2001-04-13
; NUMBER OF SEQ ID NOS: 547
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 357
; LENGTH: 241
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-834-759-357

Alignment Scores:
Pred. No.: 45.8 Length: 241
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 16.28% Indels: 0
DB: 4 Gaps: 0

US-10-624-218-3 (1-43) x US-09-834-759-357 (1-241)

Qy 15 LeuGlutThrArgValAlaGlu 21
Db 208 CTGGAGACTCGGGTGGCCGAG 188

```

Db      208 CTGGAGACTCGGTGGCCGAG 188

RESULT 11
US-09-513-999C-939
; Sequence 939, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 939
; LENGTH: 472
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 131..472
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 6
; OTHER INFORMATION: s=g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 7
; OTHER INFORMATION: n=a, g, c or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 292
; OTHER INFORMATION: y=c or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 435
; OTHER INFORMATION: n=a, g, c or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 456
; OTHER INFORMATION: s=g or c
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 102
; OTHER INFORMATION: Xaa=Ala or Glu or Gly or Val
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 109
; OTHER INFORMATION: Xaa=Ser or Thr
; US-09-513-999C-939

Alignment Scores:
Pred. No.:      85.5      Length:      472
Score:          7.00      Matches:      7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match:    16.28% Indels: 0
DB:             4      Gaps: 0

US-10-624-218-3 (1-43) x US-09-513-999C-939 (1-472)
Qy      15 LeuGluThrArgValAlaGlu 21
      |||||||
Db      23 CTGGAGACTCGGTGGCCGAG 43

RESULT 12
US-09-669-751-189
; Sequence 189, Application US/09669751

Db      208 CTGGAGACTCGGTGGCCGAG 188

RESULT 10
US-09-551-621-357/c
; Sequence 357, Application US/09551621
; Patent No. 6825175
; GENERAL INFORMATION:
; APPLICANT: Yuqiu, Jiang
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
; FILE REFERENCE: 210121.470C5
; CURRENT APPLICATION NUMBER: US/09/551,621
; CURRENT FILING DATE: 2000-04-17
; NUMBER OF SEQ ID NOS: 479
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 357
; LENGTH: 241
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-551-621-357

Alignment Scores:
Pred. No.:      45.8      Length:      241
Score:          7.00      Matches:      7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match:    16.28% Indels: 0
DB:             4      Gaps: 0

US-10-624-218-3 (1-43) x US-09-551-621-357 (1-241)
Qy      15 LeuGluThrArgValAlaGlu 21
      |||||||
Db      208 CTGGAGACTCGGTGGCCGAG 188

RESULT 9
US-09-590-751A-357/c
; Sequence 357, Application US/09590751A
; Patent No. 6756477
; GENERAL INFORMATION:
; APPLICANT: Yuqiu, Jiang
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS FOR THE THERAPY AND
; FILE REFERENCE: 210121.470C6
; CURRENT APPLICATION NUMBER: US/09/590,751A
; CURRENT FILING DATE: 2000-06-08
; NUMBER OF SEQ ID NOS: 479
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 357
; LENGTH: 241
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-590-751A-357

Alignment Scores:
Pred. No.:      45.8      Length:      241
Score:          7.00      Matches:      7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match:    16.28% Indels: 0
DB:             4      Gaps: 0

US-10-624-218-3 (1-43) x US-09-590-751A-357 (1-241)
Qy      15 LeuGluThrArgValAlaGlu 21
      |||||||
Db      208 CTGGAGACTCGGTGGCCGAG 188
```

; Patent No. 6551575
; GENERAL INFORMATION:
; APPLICANT: Greenspan, Ralph J.
; TITLE OF INVENTION: Methods for Identifying Compounds for
; TITLE OF INVENTION: Motion Sickness, Vertigo and Other Disorders Related to
; TITLE OF INVENTION: Balance and the Perception of Gravity
; FILE REFERENCE: P-NI 3864
; CURRENT APPLICATION NUMBER: US/09/669,751
; CURRENT FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US 60/168,579
; PRIOR FILING DATE: 1999-12-02
; NUMBER OF SEQ ID NOS: 261
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 189
; LENGTH: 493
; TYPE: DNA
; ORGANISM: Drosophila
US-09-669-751-189

Alignment Scores:
Pred. No.: 89 Length: 493
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 16.28% Indels: 0
DB: 4 Gaps: 0

US-10-624-218-3 (1-43) x US-09-669-751-189 (1-493)

Qy 24 GlnArgValGlnArgLeuArg 30
Db 471 CAACGTGTCCAGCGTCTCGG 491

RESULT 13

US-09-621-976-3315
; Sequence 3315, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 3315
; LENGTH: 541
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 72..302
US-09-621-976-3315

Alignment Scores:
Pred. No.: 97 Length: 541
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 16.28% Indels: 0
DB: 4 Gaps: 0

US-10-624-218-3 (1-43) x US-09-621-976-3315 (1-541)

Qy 34 SerGlnTyrArgThrArgTyr 40
Db 441 TCACAGTACAGACACAGGTAT 461

RESULT 14

US-09-513-999C-1569
; Sequence 1569, Application US/09513999C

; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 1569
; LENGTH: 548
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 153..548
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 12
; OTHER INFORMATION: s=g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 13
; OTHER INFORMATION: s=g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 136
; OTHER INFORMATION: s=g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 173
; OTHER INFORMATION: w=a or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 196
; OTHER INFORMATION: m=a or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 197
; OTHER INFORMATION: m=a or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 236
; OTHER INFORMATION: m=a or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 237
; OTHER INFORMATION: m=a or c
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 7
; OTHER INFORMATION: Xaa=Asp or Glu
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 15
; OTHER INFORMATION: Xaa=Ala or Asp or Glu
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 28
; OTHER INFORMATION: Xaa=Asp or Glu
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 29
; OTHER INFORMATION: Xaa=Pro or Thr
US-09-513-999C-1569

Alignment Scores:
Pred. No.: 98.2 Length: 548

Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 16.28% Indels: 0
 DB: 4 Gaps: 0

US-10-624-218-3 (1-43) x US-09-513-999C-1569 (1-548)

Qy 4 GluAlaAlaPheLeuGluArg 10
 Db 21 GAGCGCGCATTCCTGGAGCGC 41

RESULT 15

US-09-621-976-3034
 ; Sequence 3034, Application US/09621976
 ; Patent No. 6639063
 ; GENERAL INFORMATION:
 ; APPLICANT: Dumas Milne Edwards, J.B.
 ; APPLICANT: Jobert, S.
 ; APPLICANT: Giordano, J.Y.
 ; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
 ; FILE REFERENCE: GENSET.054PR2
 ; CURRENT APPLICATION NUMBER: US/09/621.976
 ; CURRENT FILING DATE: 2000-07-21
 ; NUMBER OF SEQ ID NOS: 19335
 ; SOFTWARE: Patent.pm
 ; SEQ ID NO 3034
 ; LENGTH: 573
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 295..477
 US-09-621-976-3034

Alignment Scores:
 Pred. No.: 102 Length: 573
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 16.28% Indels: 0
 DB: 4 Gaps: 0

US-10-624-218-3 (1-43) x US-09-621-976-3034 (1-573)

Qy 15 LeuGluThrArgValAlaGlu 21
 Db 9 CTGGAGACTCGGTGGCGGAG 29

Search completed: July 27, 2005, 16:49:12
 Job time : 102.333 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 27, 2005, 11:55:09 ; Search time 546.16 Seconds
(without alignments)
509.100 Million cell updates/sec

Title: US-10-624-218-3

Perfect score: 43

Sequence: 1 LEIEAFLERNTALETRVA.....QRVORLRNRSQVTRYGPL 43

Scoring table:
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 7277826 seqs, 3233139505 residues

Word size: 1

Total number of hits satisfying chosen parameters: 14542843

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODE=frame+p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool/p/US10624218/runat_26072005_121437_3743/app_query.fasta_1.796
-DB=Published Applications NA -QFMT=fastap -SUFFIX=olip2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=oligo
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=quality -THR MIN=1
-ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10624218 @CGN 1 1 1189 @runat_26072005_121437_3743
-NCPU=6 -ICPU=3 -NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

Published Applications NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
19: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq.*
20: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq.*
21: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq.*
22: /cgn2_6/ptodata/1/pubpna/US10J_NEW_PUB.seq.*
23: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq.*
24: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
25: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
26: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8	18.6	279	19	US-10-437-963-91075 Sequence 91075, A
2	8	18.6	564	20	US-10-425-115-131790 Sequence 131790, A
3	8	18.6	611	20	US-10-425-115-153958 Sequence 153958, A
4	8	18.6	744	19	US-10-437-963-101792 Sequence 101792, A
5	8	18.6	798	19	US-10-767-701-1863 Sequence 1863, Ap
6	8	18.6	862	18	US-10-424-599-23118 Sequence 23118, A
7	8	18.6	1273	17	US-10-260-238-778 Sequence 778, App
8	8	18.6	1427	18	US-10-425-114-31844 Sequence 31844, A
9	8	18.6	1431	18	US-10-425-114-23276 Sequence 23276, A
10	8	18.6	1599	19	US-10-437-963-45910 Sequence 45910, A
11	8	18.6	1688	18	US-10-425-114-33218 Sequence 33218, A
12	8	18.6	1853	19	US-10-437-963-3100 Sequence 3100, Ap
13	8	18.6	1872	19	US-10-437-963-6899 Sequence 6899, Ap
14	8	18.6	1962	19	US-10-437-963-70714 Sequence 70714, A
15	8	18.6	1962	20	US-10-425-115-131791 Sequence 131791, A
16	8	18.6	2004	19	US-10-437-963-66162 Sequence 66162, A
17	8	18.6	2706	19	US-10-437-963-101755 Sequence 101755, A
18	8	18.6	2880	19	US-10-437-963-1089 Sequence 1089, Ap
19	8	18.6	3099	19	US-10-437-963-101833 Sequence 101833, A
20	8	18.6	3225	19	US-10-437-963-1134 Sequence 1134, Ap
21	8	18.6	3285	19	US-10-437-963-101745 Sequence 101745, A
22	8	18.6	3489	9	US-10-437-963-82600 Sequence 82600, A
23	8	18.6	3489	9	US-09-815-242-7731 Sequence 7731, Ap
24	8	18.6	3489	17	US-10-282-122A-30110 Sequence 30110, A
25	8	18.6	3540	19	US-10-437-963-64932 Sequence 64932, A
26	8	18.6	3552	19	US-10-437-963-68425 Sequence 68425, A
27	8	18.6	3722	19	US-10-437-963-1132 Sequence 1132, Ap
28	8	18.6	3732	19	US-10-437-963-40947 Sequence 40947, A
29	8	18.6	3816	19	US-10-437-963-67487 Sequence 67487, A
30	8	18.6	3867	19	US-10-437-963-101681 Sequence 101681, A
31	8	18.6	3930	19	US-10-437-963-6618 Sequence 6618, Ap
32	8	18.6	4149	17	US-10-259-194A-309 Sequence 309, App
33	8	18.6	4152	19	US-10-437-963-68038 Sequence 68038, A
34	8	18.6	4173	19	US-10-437-963-67609 Sequence 67609, A
35	8	18.6	4266	19	US-10-437-963-101954 Sequence 101954, A
36	8	18.6	4266	19	US-10-437-963-37899 Sequence 37899, A
37	8	18.6	4347	19	US-10-437-963-69065 Sequence 69065, A
38	8	18.6	4446	19	US-10-437-963-84296 Sequence 84296, A
39	8	18.6	4482	19	US-10-437-963-41023 Sequence 41023, A
40	8	18.6	4521	19	US-10-437-963-101718 Sequence 101718, A
41	8	18.6	4590	19	US-10-437-963-101911 Sequence 101911, A
42	8	18.6	4744	19	US-10-437-963-94514 Sequence 94514, A
43	8	18.6	4767	19	US-10-437-963-66616 Sequence 66616, A
44	8	18.6	4788	19	US-10-437-963-101883 Sequence 101883, A
45	8	18.6	4797	19	US-10-437-963-101883 Sequence 101883, A

ALIGNMENTS

RESULT 1
US-10-437-963-91075
; Sequence 91075, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B

```
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 91075
; LENGTH: 279
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_89687C.1
US-10-437-963-91075

Alignment Scores:
Pred. No.:      20.3      Length:      279
Score:          8.00      Matches:      8
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:     18.60%      Indels:      0
DB:              19       Gaps:          0

US-10-624-218-3 (1-43) x US-10-437-963-91075 (1-279)

QY      19  ValAlaGluLeuArgGlnArgVal 26
Db      147 GTCCGAGAACTGCCAACACGAGTG 170

RESULT 2
US-10-425-115-131790
; Sequence 131790, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 131790
; LENGTH: 564
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_51677C.1
US-10-425-115-131790

Alignment Scores:
Pred. No.:      37.2      Length:      564
Score:          8.00      Matches:      8
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:     18.60%      Indels:      0
DB:              20       Gaps:          0

US-10-624-218-3 (1-43) x US-10-425-115-131790 (1-564)

QY      22  LeuArgGlnArgValGlnArgLeu 29
Db      214 CTTGCCAACCGGGTTCAACGCCTA 237

RESULT 3
US-10-425-115-153958/c
; Sequence 153958, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
```

```
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 153958
; LENGTH: 611
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_7198C.1
US-10-425-115-153958

Alignment Scores:
Pred. No.:      39.9      Length:      611
Score:          8.00      Matches:      8
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:     18.60%      Indels:      0
DB:              20       Gaps:          0

US-10-624-218-3 (1-43) x US-10-425-115-153958 (1-611)

QY      6  AlapheleuGluArgGluAsnThr 13
Db      515 GCCTTTCTCGAGAGAGAAAACACA 492

RESULT 4
US-10-437-963-101792
; Sequence 101792, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 101792
; LENGTH: 744
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)-(744)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_99379C.1
US-10-437-963-101792

Alignment Scores:
Pred. No.:      47.3      Length:      744
Score:          8.00      Matches:      8
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:     18.60%      Indels:      0
DB:              19       Gaps:          0

US-10-624-218-3 (1-43) x US-10-437-963-101792 (1-744)

QY      25  ArgValGlnArgLeuArgAsnArg 32
Db      441 AGGGTTCAAGATTGAGGAACAGC 464

RESULT 5
US-10-767-701-1863
```

; Sequence 1863, Application US/10767701
; Publication No. US20040172684A1

; GENERAL INFORMATION:

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
Plants and Uses Thereof For Plant Improvement

; FILE REFERENCE: 38-21(5335)B

; CURRENT APPLICATION NUMBER: US/10767,701

; CURRENT FILING DATE: 2004-01-29

; NUMBER OF SEQ ID NOS: 63128

; SEQ ID NO 1863

; LENGTH: 798

; TYPE: DNA

; ORGANISM: Sorghum bicolor

; FEATURE:

; OTHER INFORMATION: Clone ID: SORBI-28MAY03-CLUS31630_1

US-10-767-701-1863

Alignment Scores:

Pred. No.: 50.2 Length: 798

Score: 8.00 Matches: 8

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 18.60% Indels: 0

DB: 19 Gaps: 0

US-10-624-218-3 (1-43) x US-10-767-701-1863 (1-798)

Qy 22 LeuArgGlnArgValGlnArgLeu 29

Db 240 CTTCGCAACGSGTTCAACGCCTA 263

RESULT 6

US-10-424-599-23118

; Sequence 23118, Application US/10424599

; Publication No. US20040031072A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J

; APPLICANT: Kovalic, David K

; APPLICANT: Zhou Yihua

; APPLICANT: Cao Yongwei

; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With

Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53223)B

; CURRENT APPLICATION NUMBER: US/10424,599

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 285684

; SEQ ID NO 23118

; LENGTH: 862

; TYPE: DNA

; ORGANISM: Glycine max

; FEATURE:

; NAME/KEY: unsure

; LOCATION: (1)...(862)

; OTHER INFORMATION: unsure at all n locations

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT_MRT3847_120878C.1

US-10-424-599-23118

Alignment Scores:

Pred. No.: 53.6 Length: 862

Score: 8.00 Matches: 8

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 18.60% Indels: 0

DB: 18 Gaps: 0

US-10-624-218-3 (1-43) x US-10-424-599-23118 (1-862)

Qy 22 LeuArgGlnArgValGlnArgLeu 29

Db 240 CTTCGCAACGSGTTCAACGCCTA 263

Db 797 CTTCGCAACGTTGTAACAAGGCTA 820

RESULT 7

US-10-260-238-778

; Sequence 778, Application US/10260238

; Publication No. US20040016025A1

; GENERAL INFORMATION:

; APPLICANT: Budworth, Paul R.

; APPLICANT: Moughamer, Todd G.

; APPLICANT: Briggs, Steven P.

; APPLICANT: Cooper, Bret

; APPLICANT: Glazebrook, Jane

; APPLICANT: Goff, Stephen A.

; APPLICANT: Katagiri, Fumiyaki

; APPLICANT: Kreps, Joel

; APPLICANT: Provart, Nicholas

; APPLICANT: Ricke, Darrell

; APPLICANT: Zhu, Tong

; TITLE OF INVENTION: PROMOTERS FOR REGULATION OF PLANT EXPRESSION

; FILE REFERENCE: 6011-NP

; CURRENT APPLICATION NUMBER: US/10260,238

; CURRENT FILING DATE: 2002-09-26

; PRIOR APPLICATION NUMBER: US 60/325,448

; PRIOR FILING DATE: 2001-09-26

; PRIOR APPLICATION NUMBER: US 60/325,277

; PRIOR FILING DATE: 2001-09-26

; PRIOR APPLICATION NUMBER: US 60/370,620

; PRIOR FILING DATE: 2002-04-04

; NUMBER OF SEQ ID NOS: 6077

; SEQ ID NO 778

; LENGTH: 1273

; TYPE: DNA

; ORGANISM: Oryza sativa

US-10-260-238-778

Alignment Scores:

Pred. No.: 75 Length: 1273

Score: 8.00 Matches: 8

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 18.60% Indels: 0

DB: 17 Gaps: 0

US-10-624-218-3 (1-43) x US-10-260-238-778 (1-1273)

Qy 25 ArgValGlnArgLeuArgAsnArg 32

Db 391 AGGTTCAAGATTGAGGACAGA 414

RESULT 8

US-10-425-114-31844

; Sequence 31844, Application US/10425114

; Publication No. US20040034888A1

; GENERAL INFORMATION:

; APPLICANT: Liu, Jingdong

; APPLICANT: Zhou, Yihua

; APPLICANT: Kovalic, David K.

; APPLICANT: Screen, Steven E

; APPLICANT: Tabaska, Jack E

; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53313)B

; CURRENT APPLICATION NUMBER: US/10425,114

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 73128

; SEQ ID NO 31844

; LENGTH: 1427

; TYPE: DNA

; ORGANISM: Zea mays

; FEATURE:

; OTHER INFORMATION: Clone ID: UC-ZMFLB73224C04_FLI

US-10-425-114-31844

Alignment Scores:
Pred. No.: 82.8 Length: 1427
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 18.60% Indels: 0
DB: 18 Gaps: 0

US-10-624-218-3 (1-43) x US-10-425-114-31844 (1-1427)

Qy 22 LeuArgGlnArgValGlnArgLeu 29
Db 222 CTTGCCAACGGGTTCACGCCTA 245

RESULT 9

US-10-425-114-23276
; Sequence 23276, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 23276
; LENGTH: 1431
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3595-019-D9_FLI
US-10-425-114-23276

Alignment Scores:
Pred. No.: 83 Length: 1431
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 18.60% Indels: 0
DB: 18 Gaps: 0

US-10-624-218-3 (1-43) x US-10-425-114-23276 (1-1431)

Qy 22 LeuArgGlnArgValGlnArgLeu 29
Db 224 CTTGCCAACGGGTTCACGCCTA 247

RESULT 10

US-10-437-963-45910/c
; Sequence 45910, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966

; SEQ ID NO 45910
; LENGTH: 1599
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_48829C.1
US-10-437-963-45910

Alignment Scores:
Pred. No.: 91.3 Length: 1599
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 18.60% Indels: 0
DB: 19 Gaps: 0

US-10-624-218-3 (1-43) x US-10-437-963-45910 (1-1599)

Qy 18 ArgValAlaGluLeuArgGlnArg 25
Db 240 CGTGTCGCGAGCTCCGCCAGCGG 217

RESULT 11

US-10-425-114-33218
; Sequence 33218, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 33218
; LENGTH: 1688
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-2MFLMO17076F11_FLI
US-10-425-114-33218

Alignment Scores:
Pred. No.: 95.7 Length: 1688
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 18.60% Indels: 0
DB: 18 Gaps: 0

US-10-624-218-3 (1-43) x US-10-425-114-33218 (1-1688)

Qy 22 LeuArgGlnArgValGlnArgLeu 29
Db 381 CTTGCCAACGGGTTCACGCCTA 404

RESULT 12

US-10-437-963-3100
; Sequence 3100, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad

```

; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 3100
; LENGTH: 1853
; TYPE: DNA
; ORGANISM: Oryza sativa
; NAME/KEY: unsure
; LOCATION: (1)..(1853)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_102813C.1
US-10-437-963-3100

Alignment Scores:
Pred. No.: 104 Length: 1853
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 18.60% Indels: 0
DB: 19 Gaps: 0

US-10-624-218-3 (1-43) x US-10-437-963-3100 (1-1853)

Qy 22 LeuArgGlnArgValGlnArgLeu 29
Db 281 CTAGCCCAACGGTTTCACGCGCTA 304

RESULT 13
US-10-437-963-6899
; Sequence 6899, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 6899
; LENGTH: 1872
; TYPE: DNA
; ORGANISM: Oryza sativa
; NAME/KEY: unsure
; LOCATION: (1)..(1872)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_13545C.1
US-10-437-963-6899

Alignment Scores:
Pred. No.: 105 Length: 1872
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 18.60% Indels: 0
DB: 19 Gaps: 0

US-10-624-218-3 (1-43) x US-10-437-963-6899 (1-1872)

```

```

Qy 25 ArgValGlnArgLeuArgAsnArg 32
Db 1528 AGGTTCAAGATTGAGGACAGA 1551

RESULT 14
US-10-437-963-70714
; Sequence 70714, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 70714
; LENGTH: 1962
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_71259C.1
US-10-437-963-70714

Alignment Scores:
Pred. No.: 109 Length: 1962
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 18.60% Indels: 0
DB: 19 Gaps: 0

US-10-624-218-3 (1-43) x US-10-437-963-70714 (1-1962)

Qy 25 ArgValGlnArgLeuArgAsnArg 32
Db 1381 AGAGTTCAGAGATTGAGAAATAGA 1404

RESULT 15
US-10-425-115-131791
; Sequence 131791, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 131791
; LENGTH: 1962
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_51678C.1
US-10-425-115-131791

Alignment Scores:
Pred. No.: 109 Length: 1962
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0

```

Best Local Similarity: 100.00% Mismatches: 0
Query Match: 18.60% Indels: 0
DB: 20 Gaps: 0

US-10-624-218-3 (1-43) x US-10-425-115-131791 (1-1962)

Qy 22 LeuArgGlnArgValGlnArgLeu 29

Db 600 CTTGCCCAACGGGTTCACGCCTA 623

Search completed: July 27, 2005, 17:17:50
Job time : 550.16 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 27, 2005, 11:14:04 ; Search time 2989.69 Seconds

(without alignments)
547.469 Million cell updates/sec

Title: US-10-624-218-3

Perfect score: 43

Sequence: 1 LEIEAFLERNTALETRVA.....QRVQLNRNVSYTRYGPL 43

Scoring table:

OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Word size: 1

Total number of hits satisfying chosen parameters: 68475752

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:
-MODE=frame+2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US10624218/runat_26072005_121436_3698/app_query.fasta_1.796
-DB=EST -QFMT=fastap -SUFFIX=olip2n.rst -MINMATCH=0.1 -DOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10624218 @C@N_1_1_9525 @runat_26072005_121436_3698 -NCPU=6 -ICPU=3
-NO_WMAP -LARGESQUERY -NEG_SCORES=0 -WAIT_DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : EST:*

1: gb_est1: *
2: gb_est2: *
3: gb_hic: *
4: gb_est3: *
5: gb_est4: *
6: gb_est5: *
7: gb_est6: *
8: gb_gse1: *
9: gb_gse2: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	11	25.6	594	5	EX889651
2	10	23.3	749	4	BI409790
3	9	20.9	282	6	CD272139
C	9	20.9	716	2	AW510212
	9	20.9	869	9	CL832621
	9	20.9	128	2	AW216748
6	8	18.6	154	6	CA672235
C	7	8	18.6	6	CA670617
C	8	8	18.6	6	CA670617
C	9	8	18.6	5	BQ905661

C	10	8	18.6	204	6	CA693514
C	11	8	18.6	207	6	CA737713
C	12	8	18.6	214	6	CA746928
C	13	8	18.6	221	4	BJ284497
C	14	8	18.6	222	6	CA630932
C	15	8	18.6	224	4	BM137598
C	16	8	18.6	227	6	CA647556
C	17	8	18.6	231	6	CA670440
C	18	8	18.6	243	6	CA609637
C	19	8	18.6	245	6	CA736210
C	20	8	18.6	250	6	CA735516
C	21	8	18.6	252	6	CA746916
C	22	8	18.6	252	6	CA747012
C	23	8	18.6	253	6	CA670508
C	24	8	18.6	254	2	BF933999
C	25	8	18.6	258	6	CA746117
C	26	8	18.6	262	6	CA745721
C	27	8	18.6	263	6	CD800063
C	28	8	18.6	264	6	CA735513
C	29	8	18.6	264	6	CA736048
C	30	8	18.6	266	6	CA692018
C	31	8	18.6	267	9	CL965894
C	32	8	18.6	270	1	AL927695
C	33	8	18.6	270	4	BJ283461
C	34	8	18.6	271	4	BJ283665
C	35	8	18.6	272	2	BE587169
C	36	8	18.6	277	6	CA644305
C	37	8	18.6	277	6	CA745517
C	38	8	18.6	282	6	CA697902
C	39	8	18.6	283	6	CA747083
C	40	8	18.6	284	6	C71866
C	41	8	18.6	287	6	CA664035
C	42	8	18.6	288	2	BB279039
C	43	8	18.6	288	4	BJ279514
C	44	8	18.6	290	6	CA629309
C	45	8	18.6	294	6	CA676532

ALIGNMENTS

RESULT 1

EX889651

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

EX889651 594 bp mRNA linear EST 27-JUL-2004
EX889651 tcbk Oncorhynchus mykiss cDNA clone tcbk0051c.h.15 5prim,
mRNA sequence.

EX889651.2 GI:43410570

EST.

Oncorhynchus mykiss (rainbow trout)

Oncorhynchus mykiss

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Actinopterygii; Neopterygii; Teleostei; Euteleostei;

Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.

1 (bases 1 to 594)

Govouron, M., Guiguen, Y. and Le Gac, F.

Construction and primary characterization of normalized cDNA

libraries in rainbow trout, Oncorhynchus mykiss

Unpublished (2003)

On Dec 18, 2003 this sequence version replaced gi:40140051.

Contact: Guiguen Y

INRA - SCRIBE

Campus de beaulieu, RENNES cedex, 35042, France

Tel: 02.23.48.50.09

Fax: 02.23.48.50.20

Email: Yann.Guiguen@beaulieu.rennes.inra.fr

Sequence cleaned of vector, adaptor and repetitions. Contact us

at signenasupport@jouy.inra.fr to obtain the chromatogram of this

sequence.

Plate: 0051 row: h column: 15

Seq primer: ML3R.

Location/Qualifiers

1..594

/organism="Oncorhynchus mykiss"

source

FEATURES

```

/mol_type="mRNA"
/db_xref="taxon:8022"
/clone="tcbk0051c.h.15"
/tissue_type="multi-tissues"
/dev_stage="from embryos to adults"
/lab_host="DH10B"
/clone_lib="tcbk"
/notes="Vector: pT7T3D-pac; AGENAE Rainbow trout
multi-tissues - normalized + 2 subtractions; Clone
distribution : AGENAE Resource centre. Francois PIUMI,
Francois.Piumi@jouy.inra.fr, INRA, CEA Radiobiologie et
Etude du genome (LREG), Domaine de Vilvert, 78352,
Jouy-en-Josas cedex, FRANCE, +33 (0) 1.34.65.28.02, +33
(0) 1.34.65.22.73"

ORIGIN
Alignment Scores:
Pred. No.: 0.0981 Length: 749
Score: 11.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 25.58% Indels: 0
DB: 5 Gaps: 0

US-10-624-218-3 (1-43) x BX889651 (1-594)

Qy 5 ALaAlaPheLeuGluArgGluAenThrAlaLeu 15
Db 331 CGCGCATTCCTGGAGCGGAGAACACGCGCTA 363

RESULT 2
BI409790 749 bp mRNA linear EST 14-AUG-2001
LOCUS BI409790
DEFINITION 60296295AF1 NCI_CGAP_Lu33 Mus musculus cDNA clone IMAGE:5118493 5',
mRNA sequence.
ACCESSION BI409790
VERSION BI409790.1 GI:15170713
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 749)
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE NIH-MGC http://mgi.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLaM11288 row: 1 column: 14
High quality sequence start: 25
High quality sequence stop: 745.
Location/Qualifiers
1..749
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CZECH II"
/db_xref="taxon:10090"
/clone="IMAGE:5118493"
/tissue_type="pooled lung tumors"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NCI CGAP Lu33"
/notes="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a
modified polylinker; Site 1: NotI; Site 2: EcoRI; 1st
strand cDNA was prepared from mRNA obtained from pooled
lung tumors with a Not I - oligo(dT) primer [5'

```

```

TGTACCAATCTGAAGTGGAGCGGCCCTCTGTTTTTTTTTTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not
I and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo. "

```

ORIGIN

```

Alignment Scores:
Pred. No.: 1.51 Length: 749
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 23.26% Indels: 0
DB: 4 Gaps: 0

```

US-10-624-218-3 (1-43) x BI409790 (1-749)

```

Qy 18 ArgValAlaGluLeuArgGlnArgValGln 27
Db 269 CGGGTGGCTGAACCTGCGCAACGGGTACAG 298

```

RESULT 3

```

CD272139 282 bp mRNA linear EST 01-SEP-2003
LOCUS CD272139
DEFINITION T143A02369F (PHIG:A) Ectomycorrhiza plate culture Betula
pendula/Paxillus involutus mixed EST library cDNA 5', mRNA
sequence.
ACCESSION CD272139
VERSION CD272139.1 GI:34384185
KEYWORDS EST.
SOURCE Betula pendula/Paxillus involutus mixed EST library
ORGANISM Betula pendula/Paxillus involutus mixed EST library
REFERENCE 1 (bases 1 to 282)
AUTHORS Eukaryota; mixed EST libraries.
TITLE Johansson, T., Le Quere, A., Ahren, D., Soderstrom, B., Erlandsson, R.,
Lundberg, J., Uhlen, M. and Tunlid, A.
JOURNAL Transcriptional responses of Paxillus involutus and Betula pendula
during formation of ectomycorrhizal root tissue
COMMENT Mol. Plant Microbe Interact. 17 (2), 202-215 (2004)
Contact: Johansson, T.
Fungal-Host Interaction Group (FHIG)
Microbial Ecology, Institution of Ecology
Ecology Building, Lund University, SE-223 62 Lund, Sweden
Tel: +46 46 222 45 49
Fax: +46 46 222 41 58
Email: tomas.johansson@bioekol.lu.se
PCR Primers
FORWARD: P104 (5'-GGGAGCGCGCCATTGTGTT-3')
BACKWARD: P105 (5'-AGTGAGCTCGAATGCGGCC-3')
Seq primer: P104
High quality sequence stop: 282.
Location/Qualifiers
1..282
/organism="Betula pendula/Paxillus involutus mixed EST
library"
/mol_type="mRNA"
/strain="ATCC200175 (P.involutus);Skuleskogen, Skogsforsk,Sw
eden (B.pendula)"
/db_xref="taxon:231415"
/tissue_type="Ectomycorrhizal root tissue"
/dev_stage="25 days of growth after transfer/synthesis"
/lab_host="Escherichia coli BM25.8"
/clone_lib="(PHIG:A) Ectomycorrhiza plate culture"
/notes="Vector: pTriplex2; Site 1: SfiI; Site 2: SfiI; This
EST clone is originating from one of three cDNA libraries,
constructed for transcript profiling of the mycorrhizal
interaction between the basidiomycete Paxillus involutus
and Betula pendula (birch). One library represents the
developed and functional mycorrhizal root tissue
(' (PHIG:A) Ectomycorrhiza plate culture'), a second
library represents axenically grown fungus (' (PHIG:B)
Axenic plate culture') and a third library represents

```

FEATURES

```

source
1..282
/organism="Betula pendula/Paxillus involutus mixed EST
library"
/mol_type="mRNA"
/strain="ATCC200175 (P.involutus);Skuleskogen, Skogsforsk,Sw
eden (B.pendula)"
/db_xref="taxon:231415"
/tissue_type="Ectomycorrhizal root tissue"
/dev_stage="25 days of growth after transfer/synthesis"
/lab_host="Escherichia coli BM25.8"
/clone_lib="(PHIG:A) Ectomycorrhiza plate culture"
/notes="Vector: pTriplex2; Site 1: SfiI; Site 2: SfiI; This
EST clone is originating from one of three cDNA libraries,
constructed for transcript profiling of the mycorrhizal
interaction between the basidiomycete Paxillus involutus
and Betula pendula (birch). One library represents the
developed and functional mycorrhizal root tissue
(' (PHIG:A) Ectomycorrhiza plate culture'), a second
library represents axenically grown fungus (' (PHIG:B)
Axenic plate culture') and a third library represents

```


axenically grown plants ('(FHIG:C) Axenic plate culture'). Libraries were analyzed in parallel and 3555 (FHIG:A), 3964 (FHIG:B), and 2532 (FHIG:C) high-quality (PHRED 20) ESTs of >99bp have been deposited. The cDNA libraries were constructed from total RNA using the SMART cDNA library construction kit (#K1051-1, Clontech, Palo Alto, CA, USA) according to the manufacturer's instructions. Full-length cDNAs were trimmed by Sfil, fractionated and directionally ligated into (lambda)Triplex2 arms. The lambda library was converted to a plasmid library via site-specific recombination at loxp sites in a Cre+ strain (E. coli BM25.8). Plasmid clones were randomly collected and analysed by DNA sequencing using a plasmid-specific forward primer (P104)."

ORIGIN

Alignment Scores: 7.56 Length: 282
 Pred. No.: 9.00 Matches: 9
 Score: 100.00% Conservative: 0
 Percent Similarity: 100.00% Mismatches: 0
 Best Local Similarity: 100.00% Indels: 0
 Query Match: 20.93% Gaps: 0
 DB: 6

US-10-624-218-3 (1-43) x CD272139 (1-282)

Qy 25 ArgValGlnArgLeuArgAsnArgVal 33
 |||||
 Db 228 AGGTGTCAGCGTGTGAGAAATCGATG 254
 |||||

RESULT 4

AW510212/c AW510212 716 bp mRNA linear EST 03-MAR-2000
 LOCUS fwk09e02.x1 zebrafish fin day3 regeneration Danio rerio cDNA 3'
 DEFINITION similar to SW:10N3_CARAU P18520 INTERMEDIATE FILAMENT PROTEIN ON3.
 ;, mRNA sequence.

ACCESSION AW510212

VERSION AW510212.1 GI:7148290

KEYWORDS EST.

SOURCE Danio rerio (zebrafish)

ORGANISM

Danio rerio
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
 Cypriniformes; Cyprinidae; Danio.

REFERENCE

AUTHORS Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Marra,M.,
 Eddy,S., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
 Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y.,
 Persson,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R.,
 Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
 Waterston,R. and Wilson,R.
 WashU Zebrafish EST Project 1998

TITLE

JOURNAL

COMMENT

Unpublished (1998)
 Other ESTs: fk09e02.y1
 Contact: Stephen L. Johnson
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800

Fax: 314 286 1810

Email: zbrafish@watson.wustl.edu
 cDNA Library Preparation: Raymond Lee. cDNA Library Arrayed by:
 Matthew Clark. DNA Sequencing by: Washington University Genome
 Sequencing Center Clone distribution: Genome Systems, St. Louis,
 Missouri (web address: www.genomesystems.com) (email contact:
 info@genomesystems.com) and Research Genetics, Huntsville, Alabama
 (web address: www.resgen.com) (email contact: info@resgen.com) and
 RessourcenZentrumPrimarDatenbank, Berlin, Germany (web address:
 www.rzpd.de)

zebrafish identity (p-value greater than 1e-99) found to:

gi|3461726|gb|AI106462|AI106462 ab01h07_t3 2F adult heart library

Danio rerio

Seq primer: T7 ET from Amersham

High quality sequence stop: 457.

FEATURES

source

Location/Qualifiers

1..716
 /organism="Danio rerio"
 /mol_type="mRNA"
 /db_xref="taxon:7955"
 /sex="mixed male and female"
 /tissue_type="3 day fin regenerates"
 /lab_host="E. coli XL0LR"
 /clone_lib="zebrafish fin day3 regeneration"
 /note="Vector: pBK-CMV; Site 1: EcoRI; Site 2: XhoI; 1st
 strand cDNA primed with (GA)10ACTAGTCTCGAG(T)18, followed
 by second strand synthesis, and ligated to 5' adapter
 (5')-aattggcagag-3', 3'-gccgtgctc-5'. cDNA was cloned
 directionally (EcoRI/XhoI) into Stratagene Zap express
 lambda phage arms. Mass in vivo excision done to obtain
 inserts in pBK-CMV phagemid."

ORIGIN

Alignment Scores: 18 Length: 716
 Pred. No.: 9.00 Matches: 9
 Score: 100.00% Conservative: 0
 Percent Similarity: 100.00% Mismatches: 0
 Best Local Similarity: 100.00% Indels: 0
 Query Match: 20.93% Gaps: 0
 DB: 2

US-10-624-218-3 (1-43) x AWS10212 (1-716)

Qy 22 LeuArgGlnArgValGlnArgLeuArg 30
 |||||

Db 645 CTCAGGACGGGTACAGCGACTCCGT 619
 |||||

RESULT 5

CL832621 CL832621 869 bp DNA linear GSS 09-AUG-2004
 LOCUS OR_CBA0055G16.f OR_CBA Oryza rufipogon genomic clone OR_CBA0055G16
 DEFINITION 5', genomic survey sequence.

ACCESSION CL832621

VERSION CL832621.1 GI:51078231

KEYWORDS GSS.

SOURCE Oryza rufipogon

ORGANISM

Oryza rufipogon
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE

AUTHORS Kim,H., Yu,Y., Wissotski,M., Yost,D., Stum,D., Rao,K., Luo,M.,
 Jetty,R., Kudrna,D., Muller,C., Hatfield,J., Soderlund,C. and
 Wing,R.

TITLE

JOURNAL

COMMENT

OMAP project
 Unpublished (2004)
 Contact: Rod A. Wing
 Arizona Genomics Institute
 University of Arizona
 Forbes Building Room 303, Tucson, AZ 85721-0036, USA
 Tel: 520 626 9595

Fax: 520 621 1259

Email: http://genome.arizona.edu

PCR Primers

FORWARD: TAA TAC GAC TCA CTA TAG GG

BACKWARD: CAC TCA TTA GGC ACC CCA

Plate: 0055 row: G column: 16

Seq primer: TAA TAC GAC TCA CTA TAG GG

Class: BAC ends.

Location/Qualifiers

source

1..869

/organism="Oryza rufipogon"

/mol_type="genomic DNA"

/db_xref="taxon:4529"

/clone="OR_CBA0055G16"

/tissue_type="young leaves"

/dev_stage="2 week old seedlings"

/lab_host="DH10B T1 phage resistant"

/clone lib="OR_CBa"
/note="vector: pAGIBAC1; Site 1: HindIII; Site 2: HindIII;
drk treated 36 hrs before harvest"

ORIGIN

Alignment Scores:
Pred. No.: 21.6 Length: 869
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 20.93% Indels: 0
DB: 9 Gaps: 0

US-10-624-218-3 (1-43) x CL832621 (1-869)

Qy 21 GluLeuArgGlnArgValGlnArgLeu 29
Db 421 GAACTTAGACAGCGAGTGCAGAGACTT 447

RESULT 6
LOCUS AW216748 128 bp mRNA linear EST 18-MAY-2001
DEFINITION EST295462 tomato callus, TAMU Lycopersicon esculentum cDNA clone
CUEC87023, mRNA sequence.

ACCESSION AW216748
VERSION AW216748.1 GI:6527622
KEYWORDS EST.

SOURCE Lycopersicon esculentum (tomato)

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.

1 (bases 1 to 128)

REFERENCE

AUTHORS Alcala,J., Vrebalov,J., White,R., Matern,A.L., Vision,T.,
Holt,I.E., Liang,F., Upton,J., Craven,M.B., Bowman,C.L., Ahn,S.,
Ronning,C.M., Fraser,C.M., Martin,G.B., Tanksley,S.D. and
Giovannoni,J.

TITLE Generation of ESTs from tomato callus tissue

JOURNAL Unpublished (1999)

COMMENT Contact: CUGI

Clemson University Genomics Institute

Clemson University

100 Jordan Hall, Clemson, SC 29634, USA

Email: <http://www.genome.clemson.edu/orders/index.html>

5 prime sequence.

FEATURES

source

1..128
Location/Qualifiers
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="CUEC87023"
/tissue_type="callus"
/dev_stage="25-40 days old"
/lab_host="Xli-Blue MRF"
/clone lib="tomato callus, TAMU"
/note="vector: pBlueScript SK(-); Site 1: EcoRI; Site 2:
XhoI; supplier: Giovannoni laboratory; cLEC - Cotyledons
of seedlings 7-10 days post-germination were excised, cut
at both ends and placed on MS medium with no selection.
Mixed callus was harvested at 25 and 40 days and included
undifferentiated masses. Tomato Callus EST Library"

ORIGIN

Alignment Scores:
Pred. No.: 45 Length: 128
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 18.60% Indels: 0
DB: 2 Gaps: 0

US-10-624-218-3 (1-43) x AW216748 (1-128)

Qy 16 GluThrArgValAlaGluLeuArg 23
Db 42 GAGACACAGGGTTGCTGAGCTGCCA 65

RESULT 7

CA672235/c

LOCUS CA672235 154 bp mRNA linear EST 24-NOV-2002

DEFINITION wlsu2.pk016.g16 wlsu2 Triticum aestivum cDNA clone wlsu2.pk016.g16

5' end, mRNA sequence.

ACCESSION CA672235

VERSION CA672235.1 GI:25250949

KEYWORDS EST.

SOURCE Triticum aestivum (bread wheat)

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Triticum.

REFERENCE

AUTHORS 1 (bases 1 to 154)
Tingey,S.V., Powell,W., Wolters,P., Dolan,M., Hainey,C., Yuan,Z.,
Miao,G., Caraher,N. and Hanafey,M.K.

TITLE Dupont Wheat cDNA Sequence

JOURNAL Unpublished (2002)

COMMENT Contact: Scott V. Tingey

Crop Genetics

E. I. Dupont de Nemours and Company

1 Innovation Way, P.O. Box 6104, Newark, DE 19714-6104, USA

Tel: 302-631-2602

Fax: 302-631-2607

Email: Scott.V.Tingey@USA.dupont.com

Seq primer: M13.

FEATURES

Location/Qualifiers
1..154
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="Stephens"
/db_xref="taxon:4565"
/clone="wlsu2.pk016.g16"
/tissue_type="leaf"
/clone lib="wlsu2"
/note="vector: PGEM-T Easy; Site 1: EcoRI; Site 2: XhoI;
Wheat (Triticum aestivum L.) wlsu2 cDNAs substracted with
WLM0 cDNAs"

ORIGIN

Alignment Scores:
Pred. No.: 53.4 Length: 154
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 18.60% Indels: 0
DB: 6 Gaps: 0

US-10-624-218-3 (1-43) x CA672235 (1-154)

Qy 7 PheLeuGluArgGluAsnThrAla 14

Db 81 TTCTCGAGAGAGAAAACACAGCG 58

RESULT 8

CA670617/c

LOCUS CA670617 164 bp mRNA linear EST 24-NOV-2002

DEFINITION wlsu1.pk027.i4 wlsu1 Triticum aestivum cDNA clone wlsu1.pk027.i4 5'

end, mRNA sequence.

ACCESSION CA670617

VERSION CA670617.1 GI:25249331

KEYWORDS EST.

SOURCE Triticum aestivum (bread wheat)

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Triticum.

REFERENCE

1 (bases 1 to 164)

AUTHORS Tingey,S.V., Powell,W., Wolters,P., Dolan,M., Hainey,C., Yuan,Z.,
TITLE Miao,G., Caraher,N. and Hanafey,M.K.
JOURNAL DuPont Wheat CDNA Sequence
COMMENT Unpublished (2002)
 Contact: Scott V. Tingey
 Crop Genetics
 E. I. DuPont de Nemours and Company
 1 Innovation Way, P.O. Box 6104, Newark, DE 19714-6104, USA
 Tel: 302-631-2602
 Fax: 302-631-2607
 Email: Scott.V.Tingey@USA.dupont.com
 Seq primer: M13.

FEATURES Location/Qualifiers
 source
 1..164

/organism="Triticum aestivum"
 /mol_type="mRNA"
 /cultivar="Stephens"
 /db_xref="taxon:4565"
 /clone="wlsui.pk027.14"
 /tissue_type="leaf"
 /clone_lib="wlsui"
 /note="Vector: pGEM-T Easy; Wheat (Triticum aestivum L.)
 WLMK8 cDNAs subtracted with WLM0 cDNAs"

ORIGIN

Alignment Scores:
 Pred. No.: 56.7 Length: 164
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 18.60% Indels: 0
 DB: 6 Gaps: 0

US-10-624-218-3 (1-43) x CA670617 (1-164)

Qy 7 PheLeuGluArgGluAsnThrAla 14
 |||||
Db 44 TTTCTCGAGAGAGAAACACACGCG 21

RESULT 9
BQ905661/c
LOCUS Ta04_AAFCECORC_Fusarium graminearum inoculated wheat heads
DEFINITION Triticum aestivum cDNA clone Ta04_16f01, mRNA sequence.
ACCESSION BQ905661
VERSION BQ905661.1 GI:22304445
KEYWORDS EST.
SOURCE Triticum aestivum (bread wheat)
ORGANISM Triticum aestivum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Pooidae; Triticeae; Triticum.
 1 (bases 1 to 174)
 Quillet,T., Dan,H., Koul,A., Tropiano,R., Chapados,J., Couroux,P.,
 De Moors,A., Harris,L.J., Hattori,J.I., Lacroix,C., Robert,L.S.,
 Singh,J.A., Spratt,D. and Tinker,N.A.
TITLE Expressed Sequence Tags from Wheat Heads 24 Hours after Spray
JOURNAL Inoculation with Fusarium graminearum (part 4)
COMMENT Unpublished (2002)
 Contact: Quillet, Therese
 Eastern Cereal and Oilseed Research Centre
 Agriculture and Agri-food Canada
 Neatby Bldg., Central Experimental Farm, Ottawa, Ontario, KIA 0C6,
 CANADA
 Tel: (613) 759-1658
 Fax: (613) 759-1701
 Email: ouellet@agr.gc.ca.

FEATURES Location/Qualifiers
 source
 1..174
 /organism="Triticum aestivum"
 /mol_type="mRNA"
 /cultivar="Horus"

/db_xref="taxon:4565"
 /clone="Ta04_16f01"
 /tissue_type="heads"
 /dev_stage="anthesis"
 /clone_lib="Ta04_AAFCECORC_Fusarium graminearum inoculate
 d wheat heads"
 /note="Vector: pGEM-T easy; Site 1: EcoRI; Site 2: EcoRI;
 Controlled chamber-grown wheat heads were spray inoculated
 at mid-anthesis with a Fusarium graminearum macroconidial
 suspension (50,000 spores/ml) and kept under intermittent
 misting for 24 hours, then collected and immediately
 frozen in liquid nitrogen."

ORIGIN

Alignment Scores:
 Pred. NO.: 59.9 Length: 174
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 18.60% Indels: 0
 DB: 5 Gaps: 0

US-10-624-218-3 (1-43) x BQ905661 (1-174)

Qy 1 LeuGluLeuGluAlaAlaPheLeu 8
 |||||
Db 27 TTGGAGATTGAAGCAGCGTTTGTG 4

RESULT 10

CA693514/c

LOCUS wlmk4.pk0004.e2 wlmk4 Triticum aestivum cDNA clone wlmk4.pk0004.e2

DEFINITION 5' end, mRNA sequence.

ACCESSION CA693514

VERSION CA693514.1 GI:25415300

KEYWORDS EST.

SOURCE Triticum aestivum (bread wheat)

ORGANISM Triticum aestivum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Pooidae; Triticeae; Triticum.
 1 (bases 1 to 204)
 Tingey,S.V., Powell,W., Wolters,P., Dolan,M., Hainey,C., Yuan,Z.,
 Miao,G., Caraher,N. and Hanafey,M.K.
TITLE DuPont Wheat CDNA Sequence
JOURNAL Unpublished (2002)
COMMENT Contact: Scott V. Tingey
 Crop Genetics
 E. I. DuPont de Nemours and Company
 1 Innovation Way, P.O. Box 6104, Newark, DE 19714-6104, USA
 Tel: 302-631-2602
 Fax: 302-631-2607
 Email: Scott.V.Tingey@USA.dupont.com
 Seq primer: M13.

FEATURES Location/Qualifiers
 source
 1..204

/organism="Triticum aestivum"
 /mol_type="mRNA"
 /cultivar="Stephens"
 /db_xref="taxon:4565"
 /clone="wlmk4.pk0004.e2"
 /tissue_type="leaf"
 /clone_lib="wlmk4"
 /note="Vector: pBluescript SK+; Site 1: EcoRI; Site 2:
 XhoI; Wheat (Triticum aestivum L.) seedlings 4 hr after
 inoculation w/ E. graminis and
 6-iodo-3-propyl-2-propyloxy-4(3H)-quinazolinone"

ORIGIN

Alignment Scores:
 Pred. NO.: 69.4 Length: 204
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0

```

Best Local Similarity: 100.00% Mismatches: 0
Query Match: 18.60% Indels: 0
DB: 6 Gaps: 0

US-10-624-218-3 (1-43) x CA693514 (1-204)
Qy 7 PheLeuGluArgGluAsnThrAla 14
    |||||
Db 167 TTCTCGAGAGAGAAACACAGCG 144

RESULT 11
CA737713 207 bp mRNA linear EST 26-NOV-2002
LOCUS wpi2s.pk004.n16 wpi2s Triticum aestivum cDNA clone wpi2s.pk004.n16
DEFINITION 5' end, mRNA sequence.
ACCESSION CA737713
VERSION CA737713.1 GI:25553311
KEYWORDS EST.
SOURCE Triticum aestivum (bread wheat)
ORGANISM Triticum aestivum
    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
    Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
    Pooideae; Triticeae; Triticum.
    1 (bases 1 to 207)
    Tissue type="leaf"
    /lab host="DH10B"
    /clone lib="wpi2s"
    /note="Vector: pGEM-T Easy; Site 1: SmaI; Riband
    (susceptible) wheat leaves infected with Septoria tritici
    strain A, 48 hours after infection, subtracted w/
    comparable uninfected leaves"
REFERENCE
AUTHORS Tingey,S.V., Powell,W., Wolters,P., Dolan,M., Hainey,C., Yuan,Z.,
Miao,G., Caraher,N. and Hanafey,M.K.
TITLE DuPont Wheat cDNA Sequence
JOURNAL Unpublished (2002)
COMMENT Contact: Scott V. Tingey
Crop Genetics
E. I. DuPont de Nemours and Company
1 Innovation Way, P.O. Box 6104, Newark, DE 19714-6104, USA
Tel: 302-631-2602
Fax: 302-631-2607
Email: Scott.V.Tingey@USA.dupont.com
Seq primer: M13.
FEATURES
source
    1..207
    /organism="Triticum aestivum"
    /mol_type="mRNA"
    /db_xref="taxon:4565"
    /clone="wpi2s.pk004.n16"
    /tissue_type="leaf"
    /lab host="DH10B"
    /clone lib="wpi2s"
    /note="Vector: pGEM-T Easy; Site 1: SmaI; Wheat, Polk
    cultivar (resistant), infected with Septoria tritici
    strain A 48 hours after infection"
ORIGIN
Alignment Scores:
Pred. No.: 70.4 Length: 207
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 18.60% Indels: 0
DB: 6 Gaps: 0

US-10-624-218-3 (1-43) x CA737713 (1-207)
Qy 6 AlaPheLeuGluArgGluAsnThr 13
    |||||
Db 35 GCCTTCTCGAGAGAGAAACACA 58

RESULT 12
CA746928/c 214 bp mRNA linear EST 26-NOV-2002
LOCUS wri2s.pk006.i24 wri2s Triticum aestivum cDNA clone wri2s.pk006.i24
DEFINITION 5' end, mRNA sequence.
ACCESSION CA746928
VERSION CA746928.1 GI:25565320
KEYWORDS EST.
SOURCE Triticum aestivum (bread wheat)
    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
    Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
    Pooideae; Triticeae; Triticum.
    1 (bases 1 to 221)
    Tissue type="leaf"
    /lab host="DH10B"
    /clone lib="wri2s"
    /note="Vector: pGEM-T Easy; Site 1: SmaI; Riband
    (susceptible) wheat leaves infected with Septoria tritici
    strain A, 48 hours after infection, subtracted w/
    comparable uninfected leaves"
ORIGIN
Alignment Scores:
Pred. No.: 72.6 Length: 214
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 18.60% Indels: 0
DB: 6 Gaps: 0

US-10-624-218-3 (1-43) x CA746928 (1-214)
Qy 7 PheLeuGluArgGluAsnThrAla 14
    |||||
Db 124 TTCTCGAGAGAGAAACACAGCG 101

RESULT 13
BJ284497 221 bp mRNA linear EST 09-APR-2002
LOCUS wri2s.pk006.i24 wri2s Triticum aestivum cDNA clone wri2s.pk006.i24
DEFINITION 5' end, mRNA sequence.
ACCESSION BJ284497
VERSION BJ284497.1 GI:20104823
KEYWORDS EST.
SOURCE Triticum aestivum (bread wheat)
ORGANISM Triticum aestivum
    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
    Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
    Pooideae; Triticeae; Triticum.
    1 (bases 1 to 221)
    Tissue type="leaf"
    /lab host="DH10B"
    /clone lib="wri2s"
    /note="Vector: pGEM-T Easy; Site 1: SmaI; Riband
    (susceptible) wheat leaves infected with Septoria tritici
    strain A, 48 hours after infection, subtracted w/
    comparable uninfected leaves"
REFERENCE
AUTHORS Ogiwara,Y. and Murai,K.
TITLE Expressed genes in Triticum aestivum
JOURNAL Unpublished (2002)
COMMENT Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@gene.nig.ac.jp.
FEATURES
source
    1..221
    /organism="Triticum aestivum"

```

```

Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Pooideae; Triticeae; Triticum.
1 (bases 1 to 214)
Tissue type="leaf"
/lab host="DH10B"
/clone lib="wri2s"
/note="Vector: pGEM-T Easy; Site 1: SmaI; Riband
(susceptible) wheat leaves infected with Septoria tritici
strain A, 48 hours after infection, subtracted w/
comparable uninfected leaves"
REFERENCE
AUTHORS Tingey,S.V., Powell,W., Wolters,P., Dolan,M., Hainey,C., Yuan,Z.,
Miao,G., Caraher,N. and Hanafey,M.K.
TITLE DuPont Wheat cDNA Sequence
JOURNAL Unpublished (2002)
COMMENT Contact: Scott V. Tingey
Crop Genetics
E. I. DuPont de Nemours and Company
1 Innovation Way, P.O. Box 6104, Newark, DE 19714-6104, USA
Tel: 302-631-2602
Fax: 302-631-2607
Email: Scott.V.Tingey@USA.dupont.com
Seq primer: T7.
FEATURES
source
    1..214
    /organism="Triticum aestivum"
    /mol_type="mRNA"
    /db_xref="taxon:4565"
    /clone="wri2s.pk006.i24"
    /tissue_type="leaf"
    /lab host="DH10B"
    /clone lib="wri2s"
    /note="Vector: pGEM-T Easy; Site 1: SmaI; Riband
    (susceptible) wheat leaves infected with Septoria tritici
    strain A, 48 hours after infection, subtracted w/
    comparable uninfected leaves"
ORIGIN
Alignment Scores:
Pred. No.: 72.6 Length: 214
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 18.60% Indels: 0
DB: 6 Gaps: 0

US-10-624-218-3 (1-43) x CA746928 (1-214)
Qy 7 PheLeuGluArgGluAsnThrAla 14
    |||||
Db 124 TTCTCGAGAGAGAAACACAGCG 101

RESULT 13
BJ284497 221 bp mRNA linear EST 09-APR-2002
LOCUS wri2s.pk006.i24 wri2s Triticum aestivum cDNA clone wri2s.pk006.i24
DEFINITION 5' end, mRNA sequence.
ACCESSION BJ284497
VERSION BJ284497.1 GI:20104823
KEYWORDS EST.
SOURCE Triticum aestivum (bread wheat)
ORGANISM Triticum aestivum
    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
    Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
    Pooideae; Triticeae; Triticum.
    1 (bases 1 to 221)
    Tissue type="leaf"
    /lab host="DH10B"
    /clone lib="wri2s"
    /note="Vector: pGEM-T Easy; Site 1: SmaI; Riband
    (susceptible) wheat leaves infected with Septoria tritici
    strain A, 48 hours after infection, subtracted w/
    comparable uninfected leaves"
REFERENCE
AUTHORS Ogiwara,Y. and Murai,K.
TITLE Expressed genes in Triticum aestivum
JOURNAL Unpublished (2002)
COMMENT Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@gene.nig.ac.jp.
FEATURES
source
    1..221
    /organism="Triticum aestivum"

```

```

/molecule="mRNA"
/db_xref="taxon:4565"
/clone="w3r3f05"
/tissue_type="root"
/dev_stages="feekes", scale 1"
/clone_lib="Y. Ogihara unpublished cDNA library, Wh_r"

ORIGIN
Alignment Scores:
Pred. No.: 74.8 Length: 221
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 18.60% Indels: 0
DB: 4 Gaps: 0

US-10-624-218-3 (1-43) x BJ284497 (1-221)

Qy 7 PheLeuGluArgGluAsnThrAla 14
|||||
Db 38 TTTCTCGAGAGAGAAACACAGCG 61

RESULT 14
CA630932/c
LOCUS
DEFINITION wleln.pk0039.c10 wleln Triticum aestivum cDNA clone
VERSION wleln.pk0039.c10 5' end, mRNA sequence.
KEYWORDS EST.
SOURCE Triticum aestivum (bread wheat)
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Pooideae; Triticeae; Triticum.
1 (bases 1 to 222)
Tingey,S.V., Powell,W., Wolters,P., Dolan,M., Hainey,C., Yuan,Z.,
Miao,G., Caraher,N. and Hanafey,M.K.
DuPont Wheat cDNA Sequence
Unpublished (2002)
Contact: Scott V. Tingey
Crop Genetics
E. I. DuPont de Nemours and Company
1 Innovation Way, P.O. Box 6104, Newark, DE 19714-6104, USA
Tel: 302-631-2602
Fax: 302-631-2607
Email: Scott.V.Tingey@USA.dupont.com
Seq primer: M13.

FEATURES
Location/Qualifiers
1..222
/organism="Triticum aestivum"
/mol_type="mRNA"
/db_xref="taxon:4565"
/clone="wleln.pk0039.c10"
/tissue_type="leaf"
/clone_lib="wleln"
/note="Vector: pBluescript SK+, Site_1: EcoRI; Site_2:
XhoI; Wheat Triticum aestivum L.) leaf 7 day old
etiolated seedling (normalized)"

ORIGIN
Alignment Scores:
Pred. No.: 75.1 Length: 222
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 18.60% Indels: 0
DB: 6 Gaps: 0

US-10-624-218-3 (1-43) x CA630932 (1-222)

Qy 6 AlaPheLeuGluArgGluAsnThr 13

```

```

Db 136 GCCTTTCTCGAGAGAGAAACACA 113
|||||
RESULT 15
BM137598/c
LOCUS
DEFINITION BM137598 224 bp mRNA linear EST 28-NOV-2001
WHE0483-0486_I04_I042S Wheat Fusarium graminearum infected spike
cDNA library Triticum aestivum cDNA clone WHE0483-0486_I04_I04,
mRNA sequence.
ACCESSION BM137598
VERSION BM137598.1 GI:17146365
KEYWORDS EST.
SOURCE Triticum aestivum (bread wheat)
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Pooideae; Triticeae; Triticum.
1 (bases 1 to 224)
Anderson,O.D., Chao,S., Han,P.S., Heinen,S., Heia,C.C., Kang,Y.,
Kruger,W.M., Lazo,G.R., Miller,S., Muehlbauer,G.J., Miller,R.,
Pritsch,C., Rausch,C.J., Seaton,C.L., Tong,J.C., Vance,C. and
Wilson,C.F.
The structure and function of the expressed portion of the wheat
genomes - Fusarium graminearum infected spike cDNA library
Unpublished (2001)
Contact: Olin Anderson
US Department of Agriculture, Agriculture Research Service, Pacific
West Area, Western Regional Research Center
800 Buchanan Street, Albany, CA 94710, USA
Tel: 5105959773
Fax: 5105959818
Email: oanderson@pw.usda.gov
Sequences have been trimmed to remove vector sequence and low
quality sequence with phred score less than 20. No effort was taken
to identify ESTs of fungal origin from this library, thus this EST
could be of wheat or fungal origin.
Seq primer: Stratagene SK primer.

FEATURES
Location/Qualifiers
1..224
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="Sumai3"
/db_xref="taxon:4565"
/clone="WHE0483-0486_I04_I04"
/tissue_type="Spike"
/dev_stages="Adult plant"
/lab_host="E. coli SOLR"
/clone_lib="Wheat Fusarium graminearum infected spike cDNA
library"
/note="Vector: Lambda Uni-ZAP XR, excised phagemid
pBluescript SK; Site_1: EcoRI; Site_2: XhoI; Plants were
grown in the greenhouse. Spikes were sprayed at anthesis
with Fusarium graminearum. Total RNA, and poly(A) RNA were
prepared and pooled from infected spike at 0, 6, 12, 24,
36 and 48 hours after inoculation, a cDNA library was
made, and the cDNA clones were in vivo excised to give
pBluescript phagemids in G. Muehlbauer lab at the
University of Minnesota (Kruger, W.M., Muehlbauer, G.J.,
Pritsch, C., Vance, C.). The cDNA library should contain
genes of both wheat and fungal pathogen origin. Plasmid
DNA preparations and DNA sequencing were performed in the
OD Anderson lab (all other authors)."

ORIGIN
Alignment Scores:
Pred. No.: 75.8 Length: 224
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 18.60% Indels: 0
DB: 4 Gaps: 0

US-10-624-218-3 (1-43) x BM137598 (1-224)

```

Qy 7 PheLeuGluArgGluAsnThrAla 14
| | | | | | | | | | | | | | | | | |
Db 205 TTTCGAGAGAGAGAAAACACAGCG 182

Search completed: July 27, 2005, 16:41:22
Job time : 3001.69 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 26, 2005, 11:54:19 ; Search time 108.694 Seconds
(without alignments)
153.004 Million cell updates/sec

Title: US-10-624-218-4
Perfect score: 43
Sequence: 1 LEIRAAFLQRNTALRTEVA.....QEVORLENSQVETRYGPL 43

Scoring table: OLIGO
Gapop 60.0, Gapext 60.0

Searched: 2105692 seqs, 386760381 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

- Database : A_Geneseq_16Dec04:*
- 1: geneseqp1980s:*
 - 2: geneseqp1990s:*
 - 3: geneseqp2000s:*
 - 4: geneseqp2001s:*
 - 5: geneseqp2002s:*
 - 6: geneseqp2003as:*
 - 7: geneseqp2003bs:*
 - 8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	43	100.0	43	8	ADI34279	Adi34279 Zipper pe
2	12	27.9	29	8	ADI34276	Adi34276 Zipper pe
3	11	25.6	261	7	ADD47845	Add47845 Rat Prote
4	11	25.6	303	7	ADD47847	Add47847 Human Pro
5	11	25.6	303	7	ADJ68701	Adj68701 Human hea
6	11	25.6	303	8	ADJ68701	Adj68701 Human mod
7	11	25.6	303	8	ABM82367	Abm82367 Tumour-as
8	9	20.9	176	7	ADJ80180	Adj80180 Novel hum
9	9	20.9	481	4	AAW78473	Aaw78473 Human pro
10	9	20.9	481	5	ABU65153	Abu65153 Human NOV
11	9	20.9	481	5	ABU65154	Abu65154 Human NOV
12	9	20.9	481	6	ADA54593	Ada54593 Human pro
13	9	20.9	481	6	ADN61957	Adn61957 Human pro
14	9	20.9	481	8	ADN61959	Adn61959 Human nov
15	9	20.9	490	4	AAW79457	Aaw79457 Human pro
16	9	20.9	491	8	ADR66700	Adr66700 Human pro
17	9	20.9	491	8	ADR66358	Adr66358 Human pro
18	9	20.9	315	94	AAW00956	Aaw00956 4heptad-F
19	7	16.3	315	3	AAG29149	Aag29149 Arabidops
20	7	16.3	335	3	AAG29148	Aag29148 Arabidops
21	7	16.3	350	7	ADR80351	Adr80351 Human MDD
22	7	16.3	350	7	ADe28647	Ad28647 Human NOV
23	7	16.3	350	7	ADM04116	Adm04116 Human pro
24	7	16.3	350	8	ADM93392	Adm93392 Human NOV
25	7	16.3	354	7	ADC33170	Adc33170 Human nov

26	7	16.3	416	4	ABB67901	Abb67901 Drosophil
27	7	16.3	425	7	ABO64499	Ab064499 Klebaieil
28	7	16.3	455	4	ABG07271	Abg07271 Novel hum
29	7	16.3	567	6	ABU43677	Abu43677 Protein e
30	7	16.3	573	4	ABB62984	Abb62984 Drosophil
31	7	16.3	719	8	ADQ91679	Adq91679 Polyketid
32	7	16.3	774	8	ADR09729	Adr09729 Human pro
33	7	16.3	947	7	ADC31520	Adc31520 Human nov
34	7	16.3	1043	8	ADM48031	Adm48031 Polypepti
35	7	16.3	1081	4	ABB58360	Abb58360 Drosophil
36	7	16.3	1143	8	ADR73930	Adr73930 Caenorhab
37	7	16.3	1143	8	ADN23858	Adn23858 Bacterial
38	7	16.3	1873	4	ABG14982	Abg14982 Novel hum
39	6	14.0	29	8	ADI34277	Adi34277 Zipper pe
40	6	14.0	31	2	AAI14214	Aay14214 IL-4 rece
41	6	14.0	31	2	AAI14215	Aay14215 IL-4 rece
42	6	14.0	31	2	AAI14216	Aay14216 IL-4 rece
43	6	14.0	32	4	AAAB61550	Aab61550 Peptide W
44	6	14.0	32	4	AAAB61557	Aab61557 Peptide W
45	6	14.0	42	2	AAAS7002	Aar57002 N-termina

ALIGNMENTS

RESULT 1
ADI34279
ID ADI34279 standard; protein; 43 AA.
XX
AC ADI34279;
XX
DT 15-APR-2004 (first entry)
XX
DE Zipper peptide #4 for cross linking adenoviral ligands.
XX
KW adenovirus vector; Cytostatic; HER2/neu; CD40; tumor necrosis factor;
KW TNF.
XX
OS Unidentified.
XX
PN WO2004009133-A1.
XX
PD 29-JAN-2004.
XX
PF 22-JUL-2003; 2003WO-US022852.
XX
PR 22-JUL-2002; 2002US-0397951P.
XX
(VECT-) VECTORLOGICS INC.
XX
Korokhov N, Mikhveva G;
XX
WPI; 2004-132871/13.
XX
Novel recombinant adenovirus having fiber protein modified by insertion
of first zipper peptide that can crosslink to second zipper peptide-
targeting ligand fusion protein, and binding between zipper peptides
targets vector to cell.
XX
Claim 2; SEQ ID NO 4; 54pp; English.
XX
The present invention relates to a targeted recombinant adenovirus
vector. The invention is useful for expressing a heterologous protein
chosen from a tumor associated antigen, HER2/neu and carcinoembryonic
antigen, in a target cell e.g., CD40 + cell such as dendritic cells, is
useful in a method of gene transfer to CD40 + cells, where the targeted
adenovirus vector mediates transfer of the gene encoding heterologous
protein to the cell such as the dendritic cell. The vector is useful in
gene therapy techniques for treatment of tumors. Multivalent interaction
or trimeric CD40L with CD40 receptors causes CD40 ligation, which then
results in enhanced survival of these cells and secretion of cytokines
such as interleukin (IL)-1, IL-6, IL-8, IL-10, IL-12, tumor necrosis
factor (TNF)- α agr γ , macrophage inflammatory protein (MIP)-1a and enzymes

CC such as matrix metalloproteinase, CD40-CD40L interaction also enhances
CC monocyte tumoricidal activity. The present sequence represents a zipper
CC peptide to be used for cross-linking targeting ligands with adenoviral
CC virions.
XX
SQ Sequence 43 AA;

Query Match 100.0%; Score 43; DB 8; Length 43;
Best Local Similarity 100.0%; Pred. No. 5.5e-37;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LEIRAFLRQNTALRTEVALEQEVORLENSQVETRYGPL 43
DB 1 LEIRAFLRQNTALRTEVALEQEVORLENSQVETRYGPL 43

RESULT 2
ADI34276
ID ADI34276 standard; protein; 29 AA.
AC ADI34276;
XX
DT 15-APR-2004 (first entry)
XX
DE Zipper peptide #1 for cross linking adenoviral ligands.
XX
KW adenovirus vector; Cytostatic; HER2/neu; CD40; tumor necrosis factor;
XX TNF.
XX Unidentified.
XX WO2004009133-A1.
XX
XX 29-JAN-2004.
XX
XX 22-JUL-2003; 2003WO-US022852.
XX
XX 22-JUL-2002; 2002US-0397951P.
XX (VECT-) VECTORLOGICS INC.
XX
XX Korokhov N, Mikheeva G;
XX WPI; 2004-132871/13.
XX
XX Novel recombinant adenovirus having fiber protein modified by insertion
XX of first zipper peptide that can crosslink to second zipper peptide-
XX targeting ligand fusion protein, and binding between zipper peptides
XX targets vector to cell.
XX
XX Claim 2; SEQ ID NO 1; 54pp; English.
XX
XX The present invention relates to a targeted recombinant adenovirus
XX vector. The invention is useful for expressing a heterologous protein
XX chosen from a tumor associated antigen, HER2/neu and carcinoembryonic
XX antigen, in a target cell e.g., CD40 + cell such as dendritic cells, is
XX useful in a method of gene transfer to CD40 + cells, where the targeted
XX adenovirus vector mediates transfer of the gene encoding heterologous
XX protein to the cell such as the dendritic cell. The vector is useful in
XX gene therapy techniques for treatment of tumors. Multivalent interaction
XX or trimeric CD40L with CD40 receptors causes CD40 ligation, which then
XX results in enhanced survival of these cells and secretion of cytokines
XX such as interleukin (IL)-1, IL-6, IL-8, IL-10, IL-12, tumor necrosis
XX factor (TNF)- α and macrophage inflammatory protein (MIP)-1 α and enzymes
XX such as matrix metalloproteinase. CD40-CD40L interaction also enhances
XX monocyte tumoricidal activity. The present sequence represents a zipper
XX peptide to be used for cross-linking targeting ligands with adenoviral
XX virions.
XX
SQ Sequence 29 AA;

Query Match 27.9%; Score 12; DB 8; Length 29;
Best Local Similarity 100.0%; Pred. No. 4.8e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 NTALRTEVAELE 23
DB 9 NTALRTEVAELE 20

RESULT 3
ADD47845
ID ADD47845 standard; protein; 261 AA.
XX
AC ADD47845;
XX
DT 02-DEC-2004 (revised)
DT 29-JAN-2004 (first entry)
XX
DE Rat Protein AAB20032, SEQ ID NO 13541.
XX
KW Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
XX
OS Rattus norvegicus.
OS Unidentified.
XX
XX WO2003016475-A2.
XX
XX 27-FEB-2003.
XX
XX 14-AUG-2002; 2002WO-US025765.
XX
XX 14-AUG-2001; 2001US-0312147P.
XX 01-NOV-2001; 2001US-0346382P.
XX 26-NOV-2001; 2001US-0333347P.
XX
XX (GEHO) GEN HOSPITAL CORP.
XX (FARB) BAYER AG.
XX
XX Woolf C, D'urso D, Befort K, Costigan M;
XX WPI; 2003-268312/26.
XX GENBANK; AAB20032.
XX
XX New composition comprising two or more isolated polypeptides, useful for
XX preparing a medicament for treating pain in an animal.
XX
XX Example 1; Page; 1017pp; English.
XX
XX The invention discloses a composition comprising two or more isolated rat
XX or human polynucleotides or a polynucleotide which represents a fragment,
XX derivative or allelic variation of the nucleic acid sequence. Also
XX claimed are a vector comprising the novel polynucleotide, a host cell
XX comprising the vector, a method for identifying a nucleotide sequence
XX which is differentially regulated in an animal subjected to pain and a
XX kit to perform the method, an array, a method for identifying an agent
XX that increases or decreases the expression of the polynucleotide sequence
XX that is differentially expressed in neuronal tissue of a first animal
XX subjected to pain, a method for identifying a compound which regularly
XX the expression of a polynucleotide sequence which is differentially
XX expressed in an animal subjected to pain, a method for identifying a
XX compound that regulates the activity of one or more of the
XX polynucleotides, a method for producing a pharmaceutical composition, a
XX method for identifying a compound or small molecule that regulates the
XX activity in an animal of one or more of the polypeptides given in the
XX specification, a method for identifying a compound useful in treating
XX pain and a pharmaceutical composition comprising the one or more
XX polypeptides or their antibodies. The polynucleotide or the compound that
XX modulates its activity is useful for preparing a medicament for treating
XX pain (e.g. spinal segmental nerve injury (SNI)) in an animal (e.g. gene
XX injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
XX therapy). The sequence presented is a rat protein (described in Table 3
XX of the specification) which is differentially expressed during pain.
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic form directly from WIPO at


```

CC ftp.wipo.int/pub/published_pct_sequences.
XX Sequence 261 AA;
SQ
  Query Match      25.6%; Score 11; DB 7; Length 261;
  Best Local Similarity 100.0%; Pred. No. 0.0041;
  Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
  OY 12 NTALRTEVAEL 22
  Db 230 NTALRTEVAEL 240
  RESULT 4
  ADD47847
  ID ADD47847 standard; protein; 303 AA.
  XX
  AC ADD47847;
  DT 02-DEC-2004 (revised)
  DT 29-JAN-2004 (first entry)
  XX
  DE Human Protein NP_003207, SEQ ID NO 13543.
  XX
  KW Human; pain; neuronal tissue; gene therapy;
  KW spinal segmental nerve injury; chronic constriction injury; CCI;
  KW spared nerve injury; SNI; Chung.
  XX
  OS Homo sapiens.
  OS Unidentified.
  XX
  PN WO2003016475-A2.
  XX
  PD 27-FEB-2003.
  XX
  PF 14-AUG-2002; 2003WO-US025765.
  XX
  PR 14-AUG-2001; 2001US-0312147P.
  PR 01-NOV-2001; 2001US-0346382P.
  PR 26-NOV-2001; 2001US-0333347P.
  XX
  PA (GEHO ) GEN HOSPITAL CORP.
  PA (FARB ) BAYER AG.
  XX
  PI Woolf C, D'urso D, Befort K, Costigan M;
  XX
  XX WPI; 2003-268312/26.
  DR GENBANK; NP_003207.
  XX
  PT New composition comprising two or more isolated polypeptides, useful for
  PT preparing a medicament for treating pain in an animal.
  XX
  XX Example 1; Page; 1017pp; English.
  PS
  CC The invention discloses a composition comprising two or more isolated rat
  CC or human polynucleotides or a polynucleotide which represents a fragment,
  CC derivative or allelic variation of the nucleic acid sequence. Also
  CC claimed are a vector comprising the novel polynucleotide, a host cell
  CC comprising the vector, a method for identifying a nucleotide sequence
  CC which is differentially regulated in an animal subjected to pain and a
  CC kit to perform the method, an array, a method for identifying an agent
  CC that increases or decreases the expression of the polynucleotide sequence
  CC that is differentially expressed in neuronal tissue of a first animal
  CC subjected to pain, a method for identifying a compound which regulates
  CC the expression of a polynucleotide sequence which is differentially
  CC expressed in an animal subjected to pain, a method for identifying a
  CC compound that regulates the activity of one or more of the
  CC polynucleotides, a method for producing a pharmaceutical composition, a
  CC method for identifying a compound or small molecule that regulates the
  CC activity in an animal of one or more of the polypeptides given in the
  CC specification, a method for identifying a compound useful in treating
  CC pain and a pharmaceutical composition comprising the one or more
  CC polypeptides or their antibodies. The polynucleotide or the compound that
  CC
  CC modulates its activity is useful for preparing a medicament for treating
  CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
  CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
  CC therapy). The sequence presented is a human protein (described in Table 3
  CC of the specification) which is differentially expressed during pain.
  CC Note: The sequence data for this patent did not form part of the printed
  CC specification, but was obtained in electronic form directly from WIPO at
  CC ftp.wipo.int/pub/published_pct_sequences.
  XX
  SQ Sequence 303 AA;
  Query Match      25.6%; Score 11; DB 7; Length 303;
  Best Local Similarity 100.0%; Pred. No. 0.0047;
  Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
  OY 12 NTALRTEVAEL 22
  Db 272 NTALRTEVAEL 282
  RESULT 5
  ADJ68701
  ID ADJ68701 standard; protein; 303 AA.
  XX
  AC ADJ68701;
  XX
  DT 06-MAY-2004 (first entry)
  XX
  DE Human heat mitochondrial protein as a therapeutic target SeqID507.
  XX
  KW mitochondrial; human; screening assay; diabetes mellitus;
  KW Huntington's disease; osteoarthritis;
  KW Leber's hereditary optic neuropathy; LHON;
  KW mitochondrial encephalopathy lactic acidosis and stroke; MELAS;
  KW myoclonic epilepsy ragged red fibre syndrome; MERRF; cancer;
  KW neuroprotective; nootropic; antidiabetic; anticonvulsant; antiarthritic;
  KW osteopathic; ophthalmological; cytostatic.
  XX
  OS Homo sapiens.
  XX
  PN WO2003087768-A2.
  XX
  PD 23-OCT-2003.
  XX
  PF 04-APR-2003; 2003WO-US010870.
  XX
  PR 12-APR-2002; 2002US-0372843P.
  PR 17-JUN-2002; 2002US-0389987P.
  PR 20-SEP-2002; 2002US-0412418P.
  XX
  PA (MITO-) MITOKOR.
  PA (BUCK-) BUCK INST AGE RES.
  XX
  PI Ghosh SS, Fahy ED, Zhang B, Gibson BW, Taylor SW, Glenn GM;
  PI Warnock DE;
  XX
  XX WPI; 2003-845369/78.
  DR
  PT Identifying a mitochondrial target for drug screening assays and for
  PT treating diseases associated with altered mitochondrial function.
  PT comprises detecting a modified polypeptide in a sample and correlating
  PT with the disease.
  XX
  PS Claim 1; SEQ ID NO 507; 180pp; English.
  XX
  CC This invention relates to novel mitochondrial targets that can be used
  CC for therapeutic intervention in treating a disease associated with
  CC altered mitochondrial function. Specifically, it refers to a method for
  CC identifying proteins of the human heart mitochondrial proteome that are
  CC useful for drug screening assays, as well as therapeutic targets. The
  CC present invention describes a method for identifying such proteins that
  CC can be used in the treatment of various diseases associated with altered
  CC mitochondrial function including diabetes mellitus, Huntington's disease,
  CC

```

CC osteoarthritis, Leber's hereditary optic neuropathy (LHON), mitochondrial
 CC encephalopathy lactic acidosis and stroke (MELAS), myoclonic epilepsy
 CC ragged red fibre syndrome (MERRF) or cancer. Accordingly, these
 CC compositions have neuroprotective, nootropic, antidiabetic,
 CC anticonvulsant, antiarthritic, osteopathic, ophthalmological and
 CC cytotatic activities. This polypeptide sequence is a human heart
 CC mitochondrial protein of the invention.

XX Sequence 303 AA;
 SQ

Query Match 25.6%; Score 11; DB 7; Length 303;
 Best Local Similarity 100.0%; Pred. No. 0.0047;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 NTALRTEVAEL 22
 |||||
 Db 272 NTALRTEVAEL 282

RESULT 6
 ADI82509
 ID ADI82509 standard; protein; 303 AA.
 AC ADI82509;
 XX
 XX 22-APR-2004 (first entry)
 DT Human modifier of p21 (MP21) protein sequence SeqID75.
 DE
 DE p21 pathway modulating agent; assay system; MP21; cytostatic;
 KW MP21 protein activity modulator; cancer; genetically modified animal;
 KW human.
 KW
 XX Homo sapiens.
 OS
 XX WO2004005486-A2.
 PN
 XX 15-JAN-2004.
 PD
 XX 09-JUL-2003; 2003WO-US021510.
 PF
 XX 10-JUL-2002; 2002US-0394795P.
 PR 07-AUG-2002; 2002US-0401739P.
 PR 16-SEP-2002; 2002US-0411010P.
 PR 30-DEC-2002; 2002US-0437158P.
 XX (EXEL-) EXELIXIS INC.
 PA
 XX Francis-Lang H, Friedman L, Kidd T, Roche S, Joo DM, Lickteig K;
 XX Amundsen CD, Hai B, Zhang H, Adamkewicz JI, Hammonds GR;
 PI
 PI WPI; 2004-091358/09.
 DR N-PSDB; ADI82448.
 XX
 XX Identifying a candidate p21 pathway modulating agent, useful for treating
 PT a disease such as cancer, comprises contacting an assay system comprising
 PT a MP21 polypeptide or nucleic acid with a test agent.
 XX
 XX Example 2; SEQ ID NO 75; 392pp; English.
 PS
 XX This invention relates to a novel candidate p21 pathway modulating agent
 CC by contacting an assay system comprising an MP21 (modifier of p21)
 CC polypeptide or nucleic acid with a test agent, where in the absence of
 CC the test agent the system provides a reference activity and detecting a
 CC test agent-biased activity of the assay system. The invention may be
 CC useful for the production of compounds with a cytostatic activity through
 CC modulation of MP21 protein activity. The MP21 polypeptide or nucleic acid
 CC can be used for identifying MP21 modulating agents useful as therapeutic
 CC targets for diagnosing cancer or treating disorders associated with
 CC defective or impaired p21 and/or MP21 function. MP21 modulating agents
 CC are useful in diagnosis, therapy, for example treating cancer, and
 CC pharmaceutical development. The genetically modified animals may be used
 CC for in vivo assays to test for activity of a candidate p21 modulating

CC agent, or to further assess the role of MP21 in a p21 pathway process.
 CC The present sequence is that of a human MP21 protein which is an
 CC orthologue of a Drosophila p21 modifier and which was used in the
 CC exemplification of the invention.

XX Sequence 303 AA;
 SQ

Query Match 25.6%; Score 11; DB 8; Length 303;
 Best Local Similarity 100.0%; Pred. No. 0.0047;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 NTALRTEVAEL 22
 |||||
 Db 272 NTALRTEVAEL 282

RESULT 7
 ABM82367
 ID ABM82367 standard; protein; 303 AA.
 AC ABM82367;
 XX
 XX 18-NOV-2004 (first entry)
 DT Tumour-associated antigenic target (TAT) polypeptide PRO83291, SEQ:6078.
 DE
 DE Tumour-associated antigenic target; TAT; human; overexpression; cancer;
 XX tumour; diagnosis; cell proliferative disorder; breast cancer;
 KW colorectal cancer; lung cancer; ovarian cancer; liver cancer;
 KW central nervous system cancer; bladder cancer; pancreatic cancer;
 KW cervical cancer; melanoma; leukaemia; hybridisation probe;
 KW chromosome identification; chromosome mapping; gene mapping;
 KW gene therapy; cytostatic.
 KW
 XX Homo sapiens.
 OS
 XX WO2004030615-A2.
 PN
 XX 15-APR-2004.
 PD
 XX 29-SEP-2003; 2003WO-US028547.
 PF
 XX 02-OCT-2002; 2002US-0414971P.
 PR (GETH) GENENTECH INC.
 PA
 XX Wu TD, Zhang Z, Zhou Y;
 PI
 XX WPI; 2004-347921/32.
 DR N-PSDB; ACN40954.
 DR
 XX New tumor-associated antigenic target polypeptides and nucleic acids,
 PT useful in preparing a medicament for treating or detecting a
 PT proliferative disorder, e.g. breast, lung, colorectal, ovarian or
 PT prostate cancer or tumor.
 XX
 XX Claim 12; SEQ ID NO 6078; 7273pp; English.
 PS
 XX The invention relates to human tumour-associated antigenic target (TAT)
 CC polypeptides, and their related nucleic acids. The TAT polypeptides are
 CC overexpressed in cancer tissues compared to normal tissues, and may thus
 CC serve as effective targets for the diagnosis and treatment of cancer in
 CC mammals. The invention also relates to nucleic acid and polypeptide
 CC sequences at least 80% identical to the TAT nucleic acids and
 CC polypeptides; expression vectors and host cells comprising a TAT nucleic
 CC acid; an antibody specific for a TAT polypeptide; a peptide or organic
 CC molecule which binds to a TAT polypeptide; fusion proteins comprising a
 CC TAT polypeptide; and methods and compositions for the treatment or
 CC diagnosis of cancer in mammals. TAT polypeptides, nucleic acids,
 CC antibodies, antagonists, binding molecules and compositions are useful
 CC for diagnosing or treating a cell proliferative disorder associated with
 CC increased TAT expression, particularly cancers such as breast cancer,
 CC colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder

CC	infections. These may also be used as elements on a microarray which may
CC	monitor or measure protein-protein interactions, drug-target
CC	interactions, and gene expression profiles. The polynucleotide may also
CC	be used in chromosomal mapping and in various diagnostic assays. These
CC	are also useful in assessing the effects of exogenous compounds on the
CC	expression of nucleic acids and amino acid sequences of NAAP, in
CC	facilitating drug discovery process, and in investigating the
CC	pathogenesis of diseases or medical conditions. This sequence corresponds
CC	to one of the proteins of the inventions.
XX	
SQ	Sequence 176 AA;
	Query Match 20.9%; Score 9; DB 7; Length 176;
	Best Local Similarity 100.0%; Pred. No. 0.33;
	Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0
Qy	21 ELEQEVQRL 29
Db	20 ELEQEVQRL 28
RESULT 9	
AAW78473	
ID	AAW78473 standard; protein; 481 AA.
XX	
AC	AAW78473;
XX	
DT	06-NOV-2001 (first entry)
XX	
DE	Human protein SEQ ID NO 1135.
XX	
KW	Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW	vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW	tissue growth factor; immunomodulatory; cancer; leukaemia;
KW	nervous system disorder; arthritis; inflammation.
XX	
OS	Homo sapiens.
XX	
PN	WO200157190-A2.
XX	
PD	09-AUG-2001.
XX	
PF	05-FEB-2001; 2001WO-US004098.
XX	
PR	03-FEB-2000; 2000US-00496914.
PR	27-APR-2000; 2000US-00560875.
PR	20-JUN-2000; 2000US-00598075.
PR	19-JUL-2000; 2000US-00620325.
PR	01-SEP-2000; 2000US-00854936.
PR	15-SEP-2000; 2000US-00863561.
PR	20-OCT-2000; 2000US-00693325.
PR	30-NOV-2000; 2000US-00728422.
XX	
PA	(HYSE-) HYSEQ INC.
XX	
PI	Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;
PI	Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI	Xue AJ, Yang Y, Wejhrman T, Goodrich R;
XX	
DR	WPI; 2001-476283/51.
DR	N-FSDB; AAKS1606.
XX	
PT	Nucleic acids encoding polypeptides with cytokine-like activities, useful
PT	in diagnosis and gene therapy.
XX	
PS	Claim 20; Page 3367-3368; 6221pp; English.
XX	
CC	The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC	encoded polypeptides (AAW78323-AAW80302) that exhibit activity elating to
CC	cytokine, cell proliferation or cell differentiation or which may induce
CC	production of other cytokines in other cell populations. The
CC	polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC	peptide therapy. The polypeptides have various cytokine-like activities,

CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
CC (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the
CC sequence listing were missing at the time of publication
XX
SQ Sequence 481 AA;

Query Match 20.9%; Score 9; DB 4; Length 481;
Best Local Similarity 100.0%; Pred. No. 0.85;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 ELEQEVQRL 29
Db 20 ELEQEVQRL 28
|||||
|

RESULT 10
ABU65153
ID ABU65153 standard; protein; 481 AA.
XX AC ABU65153;
XX
XX 20-MAY-2003 (first entry)
XX
XX Human NOV80a protein.
DE
XX NOVX; cytostatic; cardiant; antiarteriosclerotic; antiaethmatic; cancer;
KW hypotensive; cardiomyopathy; bronchial asthma; gene therapy; vaccine;
KW human.
XX Homo sapiens.
OS
XX WO200272757-A2.
PN
XX 19-SEP-2002.
PD
XX
XX 08-MAR-2002; 2002WO-US006908.
PF
XX
XX 08-MAR-2001; 2001US-0274101P.
PR 08-MAR-2001; 2001US-0274194P.
PR 08-MAR-2001; 2001US-0274281P.
PR 08-MAR-2001; 2001US-0274322P.
PR 09-MAR-2001; 2001US-0274849P.
PR 12-MAR-2001; 2001US-0275233P.
PR 13-MAR-2001; 2001US-0275578P.
PR 13-MAR-2001; 2001US-0275579P.
PR 13-MAR-2001; 2001US-0275601P.
PR 14-MAR-2001; 2001US-0276000P.
PR 16-MAR-2001; 2001US-0276778P.
PR 19-MAR-2001; 2001US-0276994P.
PR 20-MAR-2001; 2001US-0277239P.
PR 20-MAR-2001; 2001US-0277321P.
PR 20-MAR-2001; 2001US-0277327P.
PR 21-MAR-2001; 2001US-0277791P.
PR 22-MAR-2001; 2001US-0277833P.
PR 23-MAR-2001; 2001US-0278152P.
PR 26-MAR-2001; 2001US-0278894P.
PR 27-MAR-2001; 2001US-0278999P.
PR 27-MAR-2001; 2001US-0279036P.
PR 28-MAR-2001; 2001US-0279344P.
PR 30-MAR-2001; 2001US-0277338P.
PR 30-MAR-2001; 2001US-0279995P.
PR 30-MAR-2001; 2001US-0280233P.
PR 02-APR-2001; 2001US-0280802P.
PR 02-APR-2001; 2001US-0280822P.
PR 04-APR-2001; 2001US-0280900P.
PR 13-APR-2001; 2001US-0281194P.
PR 30-APR-2001; 2001US-0283675P.
PR 02-MAY-2001; 2001US-0287424P.
PR 02-MAY-2001; 2001US-0288066P.

PR 03-MAY-2001; 2001US-0288342P.
PR 03-MAY-2001; 2001US-0288528P.
PR 15-MAY-2001; 2001US-0291190P.
PR 16-MAY-2001; 2001US-0291099P.
PR 16-MAY-2001; 2001US-0291240P.
PR 30-MAY-2001; 2001US-0294485P.
PR 31-MAY-2001; 2001US-0294889P.
PR 31-MAY-2001; 2001US-0294899P.
PR 18-JUN-2001; 2001US-0299027P.
PR 19-JUN-2001; 2001US-0299303P.
PR 19-JUN-2001; 2001US-0299310P.
PR 10-JUL-2001; 2001US-0304354P.
PR 31-JUL-2001; 2001US-0309198P.
PR 16-AUG-2001; 2001US-0312903P.
PR 10-SEP-2001; 2001US-0318462P.
PR 12-SEP-2001; 2001US-0318770P.
PR 27-SEP-2001; 2001US-0325430P.
PR 27-SEP-2001; 2001US-0325861P.
PR 18-OCT-2001; 2001US-0330380P.
PR 31-OCT-2001; 2001US-0335301P.
PR 14-NOV-2001; 2001US-0332172P.
PR 14-NOV-2001; 2001US-0332271P.
PR 14-NOV-2001; 2001US-0332272P.
PR 14-NOV-2001; 2001US-0333184P.
PR 14-NOV-2001; 2001US-0333272P.
PR 21-NOV-2001; 2001US-0332094P.
PR 03-DEC-2001; 2001US-0337426P.
PR 03-DEC-2001; 2001US-0338092P.
PR 04-DEC-2001; 2001US-0337185P.
PR 03-JAN-2002; 2002US-0345705P.
PR 07-MAR-2002; 2002US-00092900.
XX
XX (CURA-) CURAGEN CORP.
XX
XX Padigaru M, Spytek KA, Shenoy SG, Taupier RJ, Pena CEA, Li L;
PI Zerrhusen BD, Gusev V, Ji W, Gorman L, Miller CE, Kekuda R;
PI Patturajan M, Gangolli E, Vernet CAM, Guo X, Tchernev V;
PI Fernandes ER, Casman SJ, Malyankar UM, Gerlach V, Liu Y, Anderson D;
PI Spaderna SK, Catterton E, Burgess C, Leite M, Zhong H, Alsobrook JP;
PI Lepley DM, Rieger DK;
XX
XX WPI; 2002-723332/78.
DR N-PSDB; ABX97120.
XX
XX NOVX polypeptides and polynucleotides, useful for preventing or treating
PT a disorder associated with aberrant NOVX expression or activity e.g.,
PT cancer, hypertension, atherosclerosis, cardiomyopathy or bronchial
PT asthma.
XX
PS Claim 1; Page 325; 1103pp; English.
XX
XX This invention describes novel human NOVX polypeptides which have
CC cytosolic, cardiac, antiarteriosclerotic, antiasthmatic and hypotensive
CC activity. Pharmaceutical compositions comprising the NOVX proteins or
CC nucleic acid molecules or NOVX antibodies are useful for preventing or
CC treating a disorder associated with aberrant NOVX expression or activity
CC e.g. cancer, hypertension, atherosclerosis, cardiomyopathy or bronchial
CC asthma. The products of the invention can be used for gene therapy or in
CC a vaccine. ABU65041-ABU65218 represent the NOVX polypeptides encoded by
CC ABX97008-ABX97185
XX
SQ Sequence 481 AA;

Query Match 20.9%; Score 9; DB 5; Length 481;
Best Local Similarity 100.0%; Pred. No. 0.85;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 ELEQEVQRL 29
Db 20 ELEQEVQRL 28
|||||
|

RESULT 11

ABU65154
ID ABU65154 standard; protein; 481 AA.
XX AC ABU65154;
XX DT 20-MAY-2003 (first entry)
XX DE Human NOV80b protein.
XX KW NOVX; cytostatic; cardiant; antiarteriosclerotic; antiasthmatic; cancer;
KW hypotensive; cardiomyopathy; bronchial asthma; gene therapy; vaccine;
KW human.
XX OS Homo sapiens.
XX PN WO200272757-A2.
XX PD 19-SEP-2002.
XX PF 08-MAR-2002; 2002WO-US006908.
XX PR 08-MAR-2001; 2001US-0274101P.
PR 08-MAR-2001; 2001US-0274194P.
PR 08-MAR-2001; 2001US-0274281P.
PR 08-MAR-2001; 2001US-0274322P.
PR 09-MAR-2001; 2001US-0274849P.
PR 12-MAR-2001; 2001US-0275235P.
PR 13-MAR-2001; 2001US-0275578P.
PR 13-MAR-2001; 2001US-0275579P.
PR 13-MAR-2001; 2001US-0275601P.
PR 14-MAR-2001; 2001US-0276000P.
PR 16-MAR-2001; 2001US-0276776P.
PR 19-MAR-2001; 2001US-0276994P.
PR 20-MAR-2001; 2001US-0277239P.
PR 20-MAR-2001; 2001US-0277321P.
PR 20-MAR-2001; 2001US-0277327P.
PR 21-MAR-2001; 2001US-0277791P.
PR 22-MAR-2001; 2001US-0278338P.
PR 23-MAR-2001; 2001US-0278152P.
PR 26-MAR-2001; 2001US-0278894P.
PR 27-MAR-2001; 2001US-0278999P.
PR 27-MAR-2001; 2001US-0279036P.
PR 28-MAR-2001; 2001US-0279344P.
PR 30-MAR-2001; 2001US-0277338P.
PR 30-MAR-2001; 2001US-0279995P.
PR 30-MAR-2001; 2001US-0280233P.
PR 02-APR-2001; 2001US-0280802P.
PR 02-APR-2001; 2001US-0280822P.
PR 02-APR-2001; 2001US-0280900P.
PR 04-APR-2001; 2001US-0281194P.
PR 13-APR-2001; 2001US-0283675P.
PR 30-APR-2001; 2001US-0287424P.
PR 02-MAY-2001; 2001US-0288066P.
PR 03-MAY-2001; 2001US-0288342P.
PR 03-MAY-2001; 2001US-0288528P.
PR 15-MAY-2001; 2001US-0291190P.
PR 16-MAY-2001; 2001US-0291099P.
PR 16-MAY-2001; 2001US-0291240P.
PR 30-MAY-2001; 2001US-0294485P.
PR 31-MAY-2001; 2001US-0294889P.
PR 31-MAY-2001; 2001US-0294899P.
PR 18-JUN-2001; 2001US-0299027P.
PR 19-JUN-2001; 2001US-0299303P.
PR 19-JUN-2001; 2001US-0299310P.
PR 10-JUL-2001; 2001US-0304354P.
PR 31-JUL-2001; 2001US-0309198P.
PR 16-AUG-2001; 2001US-0312903P.
PR 10-SEP-2001; 2001US-0318462P.
PR 12-SEP-2001; 2001US-0318770P.
PR 27-SEP-2001; 2001US-0325430P.
PR 27-SEP-2001; 2001US-0325681P.
PR 18-OCT-2001; 2001US-0330380P.
PR 31-OCT-2001; 2001US-0335301P.

PR 14-NOV-2001; 2001US-0332172P.
PR 14-NOV-2001; 2001US-0332271P.
PR 14-NOV-2001; 2001US-0332272P.
PR 14-NOV-2001; 2001US-0333184P.
PR 14-NOV-2001; 2001US-0333272P.
PR 21-NOV-2001; 2001US-0332094P.
PR 03-DEC-2001; 2001US-0337426P.
PR 03-DEC-2001; 2001US-0338092P.
PR 04-DEC-2001; 2001US-0337185P.
PR 03-JAN-2002; 2002US-0345705P.
PR 07-MAR-2002; 2002US-00092900.
XX PA (CURA-) CURAGEN CORP.
XX PI Padigar M, Spytek KA, Shenoy SG, Taupier RJ, Pena CEA, Li L;
PI Zehrhusen BD, Gusev V, Ji W, Gorman L, Miller CE, Kekuda R;
PI Patturajan M, Gangolli E, Vernet CAM, Guo X, Tchernev V;
PI Fernandes ER, Casman SJ, Malyankar UM, Gerlach V, Liu Y, Anderson D;
PI Spaderna SK, Catterton E, Burgess C, Leite M, Zhong H, Alsobrook JP;
PI Lepley DM, Rieger DK;
XX DR WPI; 2002-723332/78.
DR N-PSDB; ABX97121.
XX PT NOVX polypeptides and polynucleotides, useful for preventing or treating
PT a disorder associated with aberrant NOVX expression or activity e.g.,
PT cancer, hypertension, atherosclerosis, cardiomyopathy or bronchial
PT asthma.
XX PS Claim 1; Page 326; 1103pp; English.
XX CC This invention describes novel human NOVX polypeptides which have
CC cytostatic, cardiac, antiarteriosclerotic, antiasthmatic and hypotensive
CC activity. Pharmaceutical compositions comprising the NOVX proteins or
CC nucleic acid molecules or NOVX antibodies are useful for preventing or
CC treating a disorder associated with aberrant NOVX expression or activity
CC e.g. cancer, hypertension, atherosclerosis, cardiomyopathy or bronchial
CC asthma. The products of the invention can be used for gene therapy or in
CC a vaccine. ABU65041-ABU65218 represent the NOVX polypeptides encoded by
CC ABX97008-ABX97185
XX SQ Sequence 481 AA;
Query Match 20.9%; Score 9; DB 5; Length 481;
Best Local Similarity 100.0%; Pred. No. 0.85;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 21 ELEQEVQRL 29
DB 20 ELEQEVQRL 28
RESULT 12
ADA54593
ID ADA54593 standard; protein; 481 AA.
XX AC ADA54593;
XX DT 20-NOV-2003 (first entry)
XX DE Human protein, SEQ ID 2161.
XX KW Cytostatic; Anti-inflammatory; Osteopathic; Neuroprotective; Nootropic;
KW Gene Therapy; human; secretory protein; membrane proteins; cancer;
KW inflammatory disease; osteoporosis; neurological disease.
XX OS Homo sapiens.
XX PN EP1293569-A2.
XX PD 19-MAR-2003.
XX PF 21-MAR-2002; 2002EP-00006586.

```
XX 14-SEP-2001; 2001JP-00328381.
PR 24-JAN-2002; 2002US-0350435P.
XX (HELI-) HELIX RES INST.
PA (REAS-) RES ASSOC BIOTECHNOLOGY.
XX Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;
XX WPI: 2003-395539/38.
DR N-PSDB; ADA52954.
DR
XX New polynucleotides encoding full-length polypeptides, e.g. secretory
PT and/or membrane proteins, useful for developing medicines for diseases in
PT which the gene is involved, or as target molecules for gene therapy.
XX
XX Claim 14; SEQ ID NO 2161; 205pp; English.
XX
XX The present invention relates to novel human secretory or membrane
CC proteins (ADA54072-ADA55710) and their coding sequences (ADA52433-
CC ADA54071). The coding sequences are useful in the gene therapy of
CC diseases caused by abnormalities of the proteins, e.g. cancer,
CC inflammatory diseases, osteoporosis or neurological disease.
XX
XX SQ Sequence 481 AA;
Query Match 20.9%; Score 9; DB 6; Length 481;
Best Local Similarity 100.0%; Pred. No. 0.85;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 21 ELEQEVQRL 29
Db 20 ELEQEVQRL 28
|||||
|||||

RESULT 13
ADN61957
ID ADN61957 standard; protein; 481 AA.
XX
XX ADN61957;
XX
XX 01-JUL-2004 (first entry)
XX
XX Human novel protein NOV80a.
XX
XX Human; NOVX; diabetes; obesity; infectious disease; anorexia;
KW cancer-associated cachexia; cancer; neurodegenerative disorder;
KW Alzheimer's disease; Parkinson's disease; immune disorder;
KW haematopoietic disorder; dyslipidaemia; chronic disease; SNP;
KW single nucleotide polymorphism.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FH Misc-difference 46
FT /note= "May be Ser as a result of a single nucleotide
FT polymorphism"
FT
FT Misc-difference 151
FT /note= "May be Val as a result of a single nucleotide
FT polymorphism"
FT
FT Misc-difference 219
FT /note= "May be Ala as a result of a single nucleotide
FT polymorphism"
FT
FT Misc-difference 460
FT /note= "May be Asp as a result of a single nucleotide
FT polymorphism"
XX
XX US2004043382-A1.
XX
XX 04-MAR-2004.
XX
```


PR 10-SEP-2001; 2001US-0318462P.
PR 12-SEP-2001; 2001US-0318770P.
PR 27-SEP-2001; 2001US-0325430P.
PR 27-SEP-2001; 2001US-0325681P.
PR 18-OCT-2001; 2001US-0330380P.
PR 31-OCT-2001; 2001US-0335301P.
PR 14-NOV-2001; 2001US-0332172P.
PR 14-NOV-2001; 2001US-0332271P.
PR 14-NOV-2001; 2001US-0332272P.
PR 14-NOV-2001; 2001US-0333184P.
PR 14-NOV-2001; 2001US-0333272P.
PR 21-NOV-2001; 2001US-0332094P.
PR 03-DEC-2001; 2001US-0337426P.
PR 03-DEC-2001; 2001US-0338092P.
PR 04-DEC-2001; 2001US-0337185P.
PR 03-JAN-2002; 2002US-0345705P.
XX
PA (PADI/) PADIGARU M.
PA (SPYT/) SPYTEK K A.
PA (SHEN/) SHENOY S G.
PA (TAUP/) TAUPIER R J.
PA (PENA/) PENA C E A.
PA (LILL/) LI L.
PA (ZERH/) ZERHUSEN B D.
PA (GUSE/) GUSEV V Y.
PA (JIWV/) JI W.
PA (GORM/) GORMAN L.
PA (MILL/) MILLER C E.
PA (KEKU/) KEKUDA R.
PA (PATT/) PATTURAJAN M.
PA (GANG/) GANGOLLI E A.
PA (VERN/) VERNET C A M.
PA (GUOX/) GUO X S.
PA (TCHE/) TCHERNEV V T.
PA (FERN/) FERNANDES E R.
PA (CASM/) CASMAN S J.
PA (MALY/) MALYANKAR U M.
PA (GERL/) GERLACH V.
PA (LIUY/) LIU Y.
PA (ANDE/) ANDERSON D W.
PA (SPAD/) SPADERNA S K.
PA (CATT/) CATTERTON E.
PA (LEIT/) LEITE M W.
PA (ZHON/) ZHONG H.
PA (ALSO/) ALSOBROOK J P.
PA (LEPL/) LEPLEY D M.
PA (RIEG/) RIEGER D K.
PA (BURG/) BURGESS C E.
XX
PI Padigaru M, Spytek KA, Shenoy SG, Taupier RJ, Pena CEA, Li L;
PI Zerhusen BD, Gusev VY, Ji W, Gorman L, Miller CE, Kekuda R;
PI Paturajan M, Gangolli EA, Vernet CAM, Guo XS, Tchernev VT;
PI Fernandes ER, Casman SU, Malyankar UM, Gerlach V, Liu Y;
PI Anderson DW, Spaderna SK, Catterton E, Leite MW, Zhong H;
PI Alsebrook JP, Lepley DM, Rieger DK, Burgess CE;
XX
DR WPI; 2004-225693/21.
DR N-PSDB; ADN61958.
XX
PT New NOVX polypeptides and nucleic acid molecules useful for diagnosing,
PT preventing or treating NOVX-associated disorders, e.g. cancer, diabetes,
PT infection or obesity, and in chromosome mapping, tissue typing or
PT pharmacogenomics.
XX
PS Claim 1; SEQ ID NO 228; 786pp; English.
XX
PS The invention relates to an isolated polypeptide (designated NOVX, or
CC NOV1-NOV127) comprising a sequence selected from 178 fully defined amino
CC acid sequences (and their mature forms, variants and fragments). Also
CC included are an isolated nucleic acid molecule encoding NOVX, a vector
CC comprising the nucleic acid, a cell comprising the vector, methods for
CC determining the presence or amount of the polypeptide or the nucleic acid
CC molecule in a sample, methods for determining the presence of or

CC predisposition to a disease associated with altered levels of expression
CC of the above polypeptide or nucleic acid molecule in a first mammalian
CC subject, a method for identifying an agent that binds to the above
CC polypeptide, a method for identifying a potential therapeutic agent for
CC use in the treatment of a pathology that is related to aberrant
CC expression or physiological interactions of the polypeptide, a method of
CC screening for a modulator of activity or of latency or predisposition to
CC a pathology associated with the polypeptide and a method for modulating
CC the activity of the polypeptide cited above. The composition and methods
CC are useful for diagnosing, preventing or treating diseases such as
CC diabetes, obesity, infectious diseases, anorexia, cancer-associated
CC cachexia, cancer, neurodegenerative disorders like Alzheimer's disease or
CC Parkinson's disease, immune disorders, haematopoietic disorders,
CC chylodermas, and other chronic diseases. These may also be used in
CC chromosome mapping, tissue typing, preventive medicine and
CC pharmacogenomics. The polypeptides are also useful as vaccines. The
CC present sequence represents a NOVX protein of the invention.
XX
SQ Sequence 481 AA;
Query Match 20.9%; Score 9; DB 8; Length 481;
Best Local Similarity 100.0%; Pred. No. 0.85;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 21 ELEQEVQRL 29
Db 20 ELEQEVQRL 28
|||||
RESULT 15
AAM79457
ID AAM79457 standard; protein; 490 AA.
XX
AC AAM79457;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human protein SEQ ID NO 3103.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation.
XX
OS Homo sapiens.
XX
XX WO200157190-A2.
XX
PD 09-AUG-2001.
XX
XX 05-FEB-2001; 2001WO-US004098.
XX
XX 03-FEB-2000; 2000US-00496914.
XX 27-APR-2000; 2000US-00560875.
XX 20-JUN-2000; 2000US-00598075.
XX 19-JUL-2000; 2000US-00620325.
XX 01-SEP-2000; 2000US-00654936.
XX 15-SEP-2000; 2000US-00663561.
XX 20-OCT-2000; 2000US-00693325.
XX 30-NOV-2000; 2000US-00728422.
XX
PA (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Drmanac RT, Auandi V, Zhou P, Xu C, Cao Y;
PI Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
XX
XX WPI; 2001-476283/51.
DR N-PSDB; AAK52590.
XX
XX Nucleic acids encoding polypeptides with cytokine-like activities, useful
PT in diagnosis and gene therapy.
XX

PS Claim 20; Page 253; 6221pp; English.

XX The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity relating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
CC (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the
CC sequence listing were missing at the time of publication

XX SQ Sequence 490 AA;

Query Match 20.9%; Score 9; DB 4; Length 490;
Best Local Similarity 100.0%; Pred. No. 0.87;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 ELEQEVQRL 29
| | | | | | | |
DB 29 ELEQEVQRL 37

Search completed: July 26, 2005, 12:14:09
Job time : 109.694 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 26, 2005, 12:02:34 ; Search time 28.9653 Seconds
(without alignments)
110.819 Million cell updates/sec

Title: US-10-624-218-4
Perfect score: 43
Sequence: 1 LEIRAAFLQRNTALRTEVA.....QEVORLENEVSQYTRYGPL 43

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 513545 seqs, 74649064 residues

Word size : 0

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/1/iaa/5A-COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B-COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A-COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B-COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS-COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	11	25.6	345	4	US-09-949-016-10895 Sequence 10895, A
2	9	20.9	96	4	US-09-621-976-7606 Sequence 7606, Ap
3	7	16.3	89	2	US-08-690-011A-52 Sequence 52, Appl
4	7	16.3	89	3	US-09-299-495F-64 Sequence 64, Appl
5	7	16.3	425	4	US-09-489-039A-11016 Sequence 11016, A
6	6	14.0	31	4	US-09-570-337B-2 Sequence 2, Appli
7	6	14.0	31	4	US-09-570-337B-3 Sequence 3, Appli
8	6	14.0	31	4	US-09-570-337B-4 Sequence 4, Appli
9	6	14.0	42	1	US-07-752-101A-8 Sequence 8, Appli
10	6	14.0	66	3	US-09-227-357-654 Sequence 654, App
11	6	14.0	67	4	US-09-248-796A-22740 Sequence 22740, A
12	6	14.0	84	4	US-09-543-681A-4424 Sequence 4424, Ap
13	6	14.0	87	1	US-08-062-472B-10 Sequence 10, Appli
14	6	14.0	91	4	US-09-543-681A-4345 Sequence 4345, Ap
15	6	14.0	141	4	US-09-286-981B-2 Sequence 2, Appli
16	6	14.0	148	4	US-09-902-540-13147 Sequence 13147, A
17	6	14.0	150	4	US-09-270-767-59622 Sequence 59622, A
18	6	14.0	155	4	US-09-543-681A-5102 Sequence 5102, Ap
19	6	14.0	160	4	US-09-949-016-10973 Sequence 10973, A
20	6	14.0	173	1	US-08-062-472B-6 Sequence 6, Appli
21	6	14.0	178	4	US-09-583-110-4001 Sequence 4001, Ap
22	6	14.0	185	2	US-08-933-750C-2 Sequence 2, Appli
23	6	14.0	185	3	US-08-965-689A-1 Sequence 1, Appli
24	6	14.0	185	3	US-09-234-613-2 Sequence 2, Appli
25	6	14.0	185	3	US-09-359-967-1 Sequence 1, Appli
26	6	14.0	185	4	US-09-216-430C-6 Sequence 6, Appli
27	6	14.0	195	1	US-07-752-101A-68 Sequence 68, Appli

28 6 14.0 197 2 US-08-432-871C-49 Sequence 49, Appl
29 6 14.0 197 4 US-09-270-956-49 Sequence 49, Appl
30 6 14.0 198 4 US-09-107-433-4068 Sequence 4068, Ap
31 6 14.0 228 4 US-09-902-540-13809 Sequence 13809, A
32 6 14.0 242 4 US-09-252-991A-32268 Sequence 32268, A
33 6 14.0 253 4 US-09-252-991A-29632 Sequence 29632, A
34 6 14.0 261 4 US-09-252-991A-29069 Sequence 29069, A
35 6 14.0 265 4 US-09-902-540-12829 Sequence 12829, A
36 6 14.0 276 4 US-09-328-352-5243 Sequence 5243, Ap
37 6 14.0 307 4 US-09-635-872A-9 Sequence 9, Appli
38 6 14.0 307 4 US-09-636-060C-9 Sequence 9, Appli
39 6 14.0 307 4 US-09-986-552-9 Sequence 9, Appli
40 6 14.0 307 4 US-09-986-552-9 Sequence 9, Appli
41 6 14.0 307 4 US-09-636-596C-9 Sequence 9, Appli
42 6 14.0 307 4 US-10-023-894-12 Sequence 12, Appli
43 6 14.0 307 4 US-10-306-686-9 Sequence 9, Appli
44 6 14.0 316 4 US-08-976-063E-24 Sequence 24, Appli
45 6 14.0 333 4 US-09-949-016-6128 Sequence 6128, Ap

ALIGNMENTS

RESULT 1
US-09-949-016-10895
; Sequence 10895, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10895
; LENGTH: 345
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-10895
Query Match 25.6%; Score 11; DB 4; Length 345;
Best Local Similarity 100.0%; Pred. No. 0.0011;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 12 NTALRTEVAEL 22
DB 314 NTALRTEVAEL 324
RESULT 2
US-09-621-976-7606
; Sequence 7606, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 7606
; LENGTH: 96

```
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-621-976-7606

Query Match      20.9%; Score 9; DB 4; Length 96;
Best Local Similarity 100.0%; Pred. No. 0.039;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 ELEQEVQRL 29
Db 20 ELEQEVQRL 28

RESULT 3
US-08-690-011A-52
; Sequence 52, Application US/08690011A
; Patent No. 5942433
; GENERAL INFORMATION:
; APPLICANT: VINSON, Charles R.
; APPLICANT: KRYLOV, Dmitry
; TITLE OF INVENTION: EXTENSION OF A PROTEIN-PROTEIN
; TITLE OF INVENTION: INTERACTION SURFACE TO INACTIVATE THE FUNCTION OF A
; TITLE OF INVENTION: CELLULAR PROTEIN
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30B
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/690,011A
; FILING DATE: 31-JUL-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/001,654
; FILING DATE: 31-JUL-1995
; APPLICATION NUMBER: 60/018,496
; FILING DATE: 29-MAY-1996
; NAME: Serunian, Leslie A.
; ATTORNEY/AGENT INFORMATION:
; REGISTRATION NUMBER: 35,353
; REFERENCE/DOCKET NUMBER: 2026-4199US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)758-4800
; TELEFAX: (212)751-6849
; INFORMATION FOR SEQ ID NO: 52:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 89 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-690-011A-52

Query Match      16.3%; Score 7; DB 2; Length 89;
Best Local Similarity 100.0%; Pred. No. 4.5;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 AELEQEV 26
Db 44 AELEQEV 50

RESULT 4
US-09-299-495F-64
; Sequence 64, Application US/09299495F
```

```
; Patent No. 6361968
; GENERAL INFORMATION:
; APPLICANT: VINSON, Charles R.
; APPLICANT: KRYLOV, Dmitry
; TITLE OF INVENTION: EXTENSION OF A PROTEIN-PROTEIN
; TITLE OF INVENTION: INTERACTION SURFACE TO INACTIVATE THE FUNCTION OF A
; TITLE OF INVENTION: CELLULAR PROTEIN
; NUMBER OF SEQUENCES: 64
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS WORD 97
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/299,495F
; FILING DATE: 26-Apr-1999
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/001,654
; FILING DATE: 31-JUL-1995
; APPLICATION NUMBER: 60/018,496
; FILING DATE: 29-MAY-1996
; APPLICATION NUMBER: 08/690,011
; FILING DATE: 31-JULY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Serunian, Leslie A.
; REGISTRATION NUMBER: 35,353
; REFERENCE/DOCKET NUMBER: 2026-4199US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)758-4800
; TELEFAX: (212)751-6849
; INFORMATION FOR SEQ ID NO: 64:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 89 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 64:
US-09-299-495F-64

Query Match      16.3%; Score 7; DB 3; Length 89;
Best Local Similarity 100.0%; Pred. No. 4.5;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 AELEQEV 26
Db 44 AELEQEV 50

RESULT 5
US-09-489-039A-11016
; Sequence 11016, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 11016
; LENGTH: 425
; TYPE: PRT
```

; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-11016

Query Match 16.3%; Score 7; DB 4; Length 425;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 RAAFLRQ 10
|||||
DB 87 RAAFLRQ 93

RESULT 6

US-09-570-337B-2
; Sequence 2, Application US/09570337B
; Patent No. 6685932
; GENERAL INFORMATION:
; APPLICANT: European Molecular Biology Laboratory (EMBL)
; APPLICANT: Dominiques, Helena
; APPLICANT: Cregut, David
; APPLICANT: Sebald, Walter
; APPLICANT: Oschkinat, Hartmut
; APPLICANT: Serrano, Luis
; TITLE OF INVENTION: Design of Coiled-Coil Dimer Derived Antagonists of 4-Helix Bundle
; FILE REFERENCE: 100564-00000
; CURRENT APPLICATION NUMBER: US/09/570,337B
; CURRENT FILING DATE: 2000-05-12
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Description of Artificial Sequence: synthetic peptide
; NAME/KEY: misc feature
; OTHER INFORMATION: Model 2, interleukin-4 mimetic peptide 4082 Kd micrometers
US-09-570-337B-2

Query Match 14.0%; Score 6; DB 4; Length 31;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28 RLENEV 33
|||||
DB 18 RLENEV 23

RESULT 7

US-09-570-337B-3
; Sequence 3, Application US/09570337B
; Patent No. 6685932
; GENERAL INFORMATION:
; APPLICANT: European Molecular Biology Laboratory (EMBL)
; APPLICANT: Dominiques, Helena
; APPLICANT: Cregut, David
; APPLICANT: Sebald, Walter
; APPLICANT: Oschkinat, Hartmut
; APPLICANT: Serrano, Luis
; TITLE OF INVENTION: Design of Coiled-Coil Dimer Derived Antagonists of 4-Helix Bundle
; FILE REFERENCE: 100564-00000
; CURRENT APPLICATION NUMBER: US/09/570,337B
; CURRENT FILING DATE: 2000-05-12
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:

; NAME/KEY: misc feature
; OTHER INFORMATION: Description of Artificial Sequence: synthetic peptide
; NAME/KEY: misc feature
; OTHER INFORMATION: Model 2a, interleukin-4 mimetic peptide 2160 Kd micrometers
US-09-570-337B-3

Query Match 14.0%; Score 6; DB 4; Length 31;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28 RLENEV 33
|||||
DB 18 RLENEV 23

RESULT 8

US-09-570-337B-4
; Sequence 4, Application US/09570337B
; Patent No. 6685932
; GENERAL INFORMATION:
; APPLICANT: European Molecular Biology Laboratory (EMBL)
; APPLICANT: Dominiques, Helena
; APPLICANT: Cregut, David
; APPLICANT: Sebald, Walter
; APPLICANT: Oschkinat, Hartmut
; APPLICANT: Serrano, Luis
; TITLE OF INVENTION: Design of Coiled-Coil Dimer Derived Antagonists of 4-Helix Bundle
; FILE REFERENCE: 100564-00000
; CURRENT APPLICATION NUMBER: US/09/570,337B
; CURRENT FILING DATE: 2000-05-12
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Description of Artificial Sequence: synthetic peptide
; NAME/KEY: misc feature
; OTHER INFORMATION: Model 2b, interleukin-4 mimetic peptide 193 Kd micrometers
US-09-570-337B-4

Query Match 14.0%; Score 6; DB 4; Length 31;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28 RLENEV 33
|||||
DB 18 RLENEV 23

RESULT 9

US-07-752-101A-8
; Sequence 8, Application US/07752101A
; Patent No. 5326857
; GENERAL INFORMATION:
; APPLICANT: Yamamoto, Fumi-ichiro
; APPLICANT: White, Thayer
; APPLICANT: Hakomori, Sen-itiroh
; APPLICANT: Clausen, Henrik
; TITLE OF INVENTION: ABO GENOTYPING
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed and Berry
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: U.S.
; ZIP: 98104
; COMPUTER READABLE FORM: disk
; MEDIUM TYPE: Floppy

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/752,101A
FILING DATE: 19910829
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Sharkey, Richard G.
REGISTRATION NUMBER: 32,629
REFERENCE/DOCKET NUMBER: 150036.406C1
TELEPHONE: 206-622-4900
TELEFAX: 206-682-6031
TELEX: 3723836
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 42 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
FEATURE:
NAME/KEY: Region
LOCATION: 28
OTHER INFORMATION: /label= unsure
OTHER INFORMATION: /note= "This amino acid is unknown."
US-07-752-101A-8

Query Match 14.0%; Score 6; DB 1; Length 42;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 24 QEVQRL 29
Db 17 QEVQRL 22

RESULT 10

US-09-227-357-654
Sequence 654, Application US/09227357
Patent No. 6342581
GENERAL INFORMATION:
APPLICANT: Fischer et al.
TITLE OF INVENTION: 123 Human Secreted Proteins
FILE REFERENCE: P2010P1
CURRENT APPLICATION NUMBER: US/09/227,357
CURRENT FILING DATE: 1999-01-08
EARLIER APPLICATION NUMBER: PCT/US98/13684
EARLIER FILING DATE: 1998-07-07
EARLIER APPLICATION NUMBER: 60/051,926
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,793
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,925
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,929
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,803
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,732
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,931
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,932
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,916
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,930
EARLIER FILING DATE: 1997-07-08

EARLIER APPLICATION NUMBER: 60/051,918
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,920
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,733
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,795
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,919
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,928
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/055,722
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,723
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,948
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,949
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,953
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,950
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,947
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,964
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/056,360
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,684
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,984
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,954
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/058,785
EARLIER FILING DATE: 1997-09-12
EARLIER APPLICATION NUMBER: 60/058,664
EARLIER FILING DATE: 1997-09-12
EARLIER APPLICATION NUMBER: 60/058,660
EARLIER FILING DATE: 1997-09-12
EARLIER APPLICATION NUMBER: 60/058,661
EARLIER FILING DATE: 1997-09-12
NUMBER OF SEQ ID NOS: 672
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 654
LENGTH: 66
TYPE: PRT
ORGANISM: Homo sapiens
US-09-227-357-654

Query Match 14.0%; Score 6; DB 3; Length 66;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 RAAFLR 9

Db 50 RAAFLR 55

RESULT 11

US-09-248-796A-22740
Sequence 22740, Application US/09248796A
Patent No. 6747137
GENERAL INFORMATION:
APPLICANT: Keith Weinstock et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.132
CURRENT APPLICATION NUMBER: US/09/248,796A
CURRENT FILING DATE: 1999-02-12
PRIOR APPLICATION NUMBER: US 60/074,725

; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 22740
; LENGTH: 67
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-22740

Query Match 14.0%; Score 6; DB 4; Length 67;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 FLQRN 12
| | | | |
DB 14 FLQRN 19

RESULT 12
US-09-543-681A-4424
; Sequence 4424, Application US/09543681A
; Patent No. 6605709

; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 4424
; LENGTH: 84
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-4424

Query Match 14.0%; Score 6; DB 4; Length 84;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 SOYETR 39
| | | | |
DB 69 SOYETR 74

RESULT 13
US-08-062-472B-10
; Sequence 10, Application US/08062472B
; Patent No. 569594

; GENERAL INFORMATION:
; APPLICANT: Sherwood, Nancy G M
; APPLICANT: Parker, David B
; APPLICANT: McRory, John E
; APPLICANT: Lescheid, David W
; TITLE OF INVENTION: DNA ENCODING TWO FISH NEUROPEPTIDES
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: KLARQUIST, SPARKMAN, CAMPBELL, LEIGH &
; ADDRESS: WHINSTON, LLP
; STREET: ONE WORLD TRADE CENTER, SUITE 1600, 121 S.W.
; STREET: SALMON STREET
; CITY: PORTLAND
; STATE: OREGON
; COUNTRY: USA
; ZIP: 97204-2988

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/062,472B
; FILING DATE: 14-MAY-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: POLLEY, RICHARD J
; REGISTRATION NUMBER: 28107
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (503) 226-7391
; TELEFAX: (503) 228-9446
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 87 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-062-472B-10

Query Match 14.0%; Score 6; DB 1; Length 87;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28 RLENEV 33
| | | | |
DB 33 RLENEV 38

RESULT 14
US-09-543-681A-4345
; Sequence 4345, Application US/09543681A
; Patent No. 6605709

; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 4345
; LENGTH: 91
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-4345

Query Match 14.0%; Score 6; DB 4; Length 91;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 AELEQE 25
| | | | |
DB 41 AELEQE 46

RESULT 15
US-09-286-981B-2
; Sequence 2, Application US/09286981B
; Patent No. 6503511

; GENERAL INFORMATION:
; APPLICANT: Wizemann, Theresa M.
; APPLICANT: Koenig, Scott
; APPLICANT: Johnson, Leslie S
; TITLE OF INVENTION: Derivatives of Choline Binding Proteins for Vaccines
; FILE REFERENCE: 469201-396
; CURRENT APPLICATION NUMBER: US/09/286,981B
; CURRENT FILING DATE: 1999-04-06
; PRIOR APPLICATION NUMBER: US 60/085,743
; PRIOR FILING DATE: 1998-05-15
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.1

```
; SEQ ID NO 2
; LENGTH: 141
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-286-981B-2

Query Match      14.0%; Score 6; DB 4; Length 141;
Best Local Similarity 100.0%; Pred. No. 77;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      25 EVQRLE 30
Db      37 EVQRLE 42
```

Search completed: July 26, 2005, 12:22:32
Job time : 29.9653 secs

RESULT 2

```
US-10-624-218-1
; Sequence 1, Application US/10624218
; Publication No. US20040171159A1
; GENERAL INFORMATION:
; APPLICANT: Korokhov, Nikolay
; APPLICANT: Mikheeva, Galina
; TITLE OF INVENTION: Adenoviral Vector Incorporating Zipper
; TITLE OF INVENTION: Peptide-Modified Fiber Protein and Uses Thereof
; FILE REFERENCE: D6463
; CURRENT APPLICATION NUMBER: US/10/624,218
; CURRENT FILING DATE: 2003-07-22
; PRIOR APPLICATION NUMBER: US 60/397,951
; PRIOR FILING DATE: 2002-07-22
; NUMBER OF SEQ ID NOS: 4
; SEQ ID NO 1
; LENGTH: 29
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: PEPTIDE
; OTHER INFORMATION: zipper peptide E E34
US-10-624-218-1

Query Match      27.9%; Score 12; DB 16; Length 29;
Best Local Similarity 100.0%; Pred. No. 9e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      12 NTALRTEVAELE 23
Db      9 NTALRTEVAELE 20
|||||

RESULT 3
US-10-408-765A-507
; Sequence 507, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 507
; LENGTH: 303
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-507

Query Match      25.6%; Score 11; DB 16; Length 303;
Best Local Similarity 100.0%; Pred. No. 0.0074;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      12 NTALRTEVAEL 22
Db      272 NTALRTEVAEL 282
|||||

RESULT 4
US-10-094-749-2161
; Sequence 2161, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
```

```
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: NAGAI, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHICO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
; FILE REFERENCE: 084335/0160
; CURRENT APPLICATION NUMBER: US/10/094,749
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/350,435
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2161
; LENGTH: 481
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-094-749-2161

Query Match      20.9%; Score 9; DB 15; Length 481;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      21 ELEQEVQRL 29
Db      20 ELEQEVQRL 28
|||||

RESULT 5
US-10-092-900A-226
; Sequence 226, Application US/10092900A
; Publication No. US20040043382A1
; GENERAL INFORMATION:
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Taupier Jr., Raymond J.
; APPLICANT: Pena, Carol E.A.
; APPLICANT: Li, Li
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Ji, Weizhen
; APPLICANT: Gorman, Linda
; APPLICANT: Miller, Charles E.
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Patturajan, Meera
; APPLICANT: Gangolli, Esha A.M.
; APPLICANT: Vernet, Corine A.M.
; APPLICANT: Guo, Xiaojia Sasha T.
; APPLICANT: Tchernev, Elma R.
; APPLICANT: Fernandes, Stacie J.
; APPLICANT: Casman, Uziel M.
; APPLICANT: Malyankar, Valerie
; APPLICANT: Gerlach, Valerie
; APPLICANT: Liu, Yi
; APPLICANT: Anderson, David W.
; APPLICANT: Spaderna, Steven K.
; APPLICANT: Catterton, Elina
; APPLICANT: Leite, Mario W.
; APPLICANT: Zhong, Haihong
```

; APPLICANT: Alsobrook, John P.
 ; APPLICANT: Lepley, Denise M.
 ; APPLICANT: Rieger, Daniel K.
 ; APPLICANT: Burgess, Catherine E.
 ; TITLE OF INVENTION: No. US20040043382A1el Proteins and Nucleic Acids Encoding Same
 ; FILE REFERENCE: 21402-290C
 ; CURRENT APPLICATION NUMBER: US/10/092,900A
 ; CURRENT FILING DATE: 2002-03-07
 ; PRIOR FILING DATE: 2001-03-08
 ; PRIOR APPLICATION NUMBER: USN 60/274,322
 ; PRIOR FILING DATE: 2001-04-13
 ; PRIOR APPLICATION NUMBER: USN 60/283,675
 ; PRIOR FILING DATE: 2001-12-03
 ; PRIOR APPLICATION NUMBER: USN 60/338,092
 ; PRIOR FILING DATE: 2001-03-08
 ; PRIOR APPLICATION NUMBER: USN 60/274,191
 ; PRIOR FILING DATE: 2001-09-27
 ; PRIOR APPLICATION NUMBER: USN 60/325,681
 ; PRIOR FILING DATE: 2001-07-10
 ; PRIOR APPLICATION NUMBER: USN 60/304,354
 ; PRIOR FILING DATE: 2001-05-31
 ; PRIOR APPLICATION NUMBER: USN 60/294,899
 ; PRIOR FILING DATE: 2001-04-30
 ; PRIOR APPLICATION NUMBER: USN 60/287,424
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 768
 ; SEQ ID NO 226
 ; LENGTH: 481
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-10-092-900A-228

Query Match 20.9%; Score 9; DB 15; Length 481;
 Best Local Similarity 100.0%; Pred. No. 1.2;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 ELEQEVQRL 29
 |||||
 Db 20 ELEQEVQRL 28

RESULT 6
 US-10-092-900A-228
 ; Sequence 228, Application US/10092900A
 ; Publication No. US20040043382A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Padigar, Muralidhara
 ; APPLICANT: Spytek, Kimberly A.
 ; APPLICANT: Shenoy, Suresh G.
 ; APPLICANT: Taupier Jr., Raymond J.
 ; APPLICANT: Pena, Carol E.A.
 ; APPLICANT: Li, Li
 ; APPLICANT: Zerhusen, Bryan D.
 ; APPLICANT: Gusev, Vladimir Y.
 ; APPLICANT: Ji, Weizhen
 ; APPLICANT: Gorman, Linda
 ; APPLICANT: Miller, Charles E.
 ; APPLICANT: Kekuda, Ramesh
 ; APPLICANT: Patturajan, Meera
 ; APPLICANT: Gangolli, Esha A.
 ; APPLICANT: Vernet, Corine A.M.
 ; APPLICANT: Guo, Xiaojia Sasha
 ; APPLICANT: Tchernev, Velizar T.
 ; APPLICANT: Fernandes, Elma R.
 ; APPLICANT: Casman, Stacie J.
 ; APPLICANT: Malyankar, Uriel M.
 ; APPLICANT: Gerlach, Valerie
 ; APPLICANT: Liu, Yi
 ; APPLICANT: Anderson, David W.

; APPLICANT: Spaderna, Steven K.
 ; APPLICANT: Catterton, Elina
 ; APPLICANT: Leite, Mario W.
 ; APPLICANT: Zhong, Haihong
 ; APPLICANT: Alsobrook, John P.
 ; APPLICANT: Lepley, Denise M.
 ; APPLICANT: Rieger, Daniel K.
 ; APPLICANT: Burgess, Catherine E.
 ; TITLE OF INVENTION: No. US20040043382A1el Proteins and Nucleic Acids Encoding Same
 ; FILE REFERENCE: 21402-290C
 ; CURRENT APPLICATION NUMBER: US/10/092,900A
 ; CURRENT FILING DATE: 2002-03-07
 ; PRIOR FILING DATE: 2001-03-08
 ; PRIOR APPLICATION NUMBER: USN 60/274,322
 ; PRIOR FILING DATE: 2001-04-13
 ; PRIOR APPLICATION NUMBER: USN 60/283,675
 ; PRIOR FILING DATE: 2001-12-03
 ; PRIOR APPLICATION NUMBER: USN 60/338,092
 ; PRIOR FILING DATE: 2001-03-08
 ; PRIOR APPLICATION NUMBER: USN 60/274,281
 ; PRIOR FILING DATE: 2001-09-27
 ; PRIOR APPLICATION NUMBER: USN 60/325,681
 ; PRIOR FILING DATE: 2001-07-10
 ; PRIOR APPLICATION NUMBER: USN 60/304,354
 ; PRIOR FILING DATE: 2001-05-31
 ; PRIOR APPLICATION NUMBER: USN 60/294,899
 ; PRIOR FILING DATE: 2001-04-30
 ; PRIOR APPLICATION NUMBER: USN 60/287,424
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 768
 ; SEQ ID NO 228
 ; LENGTH: 481
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-10-092-900A-228

Query Match 20.9%; Score 9; DB 15; Length 481;
 Best Local Similarity 100.0%; Pred. No. 1.2;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 ELEQEVQRL 29
 |||||
 Db 20 ELEQEVQRL 28

RESULT 7
 US-10-424-599-231982
 ; Sequence 231982, Application US/10424599
 ; Publication No. US20040031072A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa Thomas J
 ; APPLICANT: Kovalic David K
 ; APPLICANT: Zhou Yihua
 ; APPLICANT: Cao Yongwei
 ; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
 ; FILE REFERENCE: 38-21(53223)B
 ; CURRENT APPLICATION NUMBER: US/10/424,599
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 285684
 ; SEQ ID NO 231982
 ; LENGTH: 151
 ; TYPE: PRT
 ; ORGANISM: Glycine max
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: PAT_MRT3847_51501C.1.pap
 ; US-10-424-599-231982

Query Match 18.6%; Score 8; DB 15; Length 151;

Best Local Similarity 100.0%; Pred. No. 4.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 28 RLENEVSQ 35
|||||

Db 62 RLENEVSQ 69

RESULT 8
US-10-424-599-173831
; Sequence 173831, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424.599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 173831
; LENGTH: 166
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_127989C.1.pap
; US-10-424-599-173831

Query Match 18.6%; Score 8; DB 15; Length 166;
Best Local Similarity 100.0%; Pred. No. 4.5;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 LRTEVAEL 22
|||||

Db 112 LRTEVAEL 119

RESULT 9
US-10-425-115-258280
; Sequence 258280, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425.115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 258280
; LENGTH: 62
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(62)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_167136C.1.pap
; US-10-425-115-258280

Query Match 16.3%; Score 7; DB 16; Length 62;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 AAFLRQR 11
|||||

Db 7 AAFLRQR 13

RESULT 10
US-10-059-720-64
; Sequence 64, Application US/10059720
; Publication No. US20030027314A1
; GENERAL INFORMATION:
; APPLICANT: VINSON, Charles R.
; APPLICANT: KRYLOV, Dmitry
; TITLE OF INVENTION: EXTENSION OF A PROTEIN-PROTEIN
; INTERACTION SURFACE TO INACTIVATE THE FUNCTION OF A
; CELLULAR PROTEIN
; NUMBER OF SEQUENCES: 64
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS WORD 97
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/059,720
; FILING DATE: 29-Jan-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/001,654
; FILING DATE: 31-JUL-1995
; APPLICATION NUMBER: 60/018,496
; FILING DATE: 29-MAY-1996
; APPLICATION NUMBER: 08/690,011
; FILING DATE: 31-JULY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Serunian, Leslie A.
; REGISTRATION NUMBER: 35,353
; REFERENCE/DOCKET NUMBER: 2026-4199US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)758-4800
; TELEFAX: (212)751-6849
; INFORMATION FOR SEQ ID NO: 64:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 89 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 64:
US-10-059-720-64

Query Match 16.3%; Score 7; DB 14; Length 89;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 AELEQEV 26
|||||

Db 44 AELEQEV 50

RESULT 11
US-10-767-701-61877
; Sequence 61877, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53535)B

; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 61877
; LENGTH: 101
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: 9853147.pep
US-10-767-701-61877

Query Match 16.3%; Score 7; DB 16; Length 101;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 VAELEQE 25
| | | | |
Db 18 VAELEQE 24

RESULT 12
US-10-425-115-306881
; Sequence 306881, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 306881
; LENGTH: 124
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_42946C.1.pep
US-10-425-115-306881

Query Match 16.3%; Score 7; DB 16; Length 124;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 26 VORLENE 32
| | | | |
Db 105 VORLENE 111

RESULT 13
US-10-425-115-288190
; Sequence 288190, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 288190
; LENGTH: 139
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_25920C.1.pep

US-10-425-115-288190

Query Match 16.3%; Score 7; DB 16; Length 139;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 29 LENEVSQ 35
| | | | |
Db 39 LENEVSQ 45

RESULT 14
US-10-424-599-247052
; Sequence 247052, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 247052
; LENGTH: 183
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(183)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_65117C.1.pep
US-10-424-599-247052

Query Match 16.3%; Score 7; DB 15; Length 183;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 ELEQEVQ 27
| | | | |
Db 71 ELEQEVQ 77

RESULT 15
US-10-425-114-67793
; Sequence 67793, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 67793
; LENGTH: 210
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3637-221-H8_FLI.pep
US-10-425-114-67793

Query Match 16.3%; Score 7; DB 15; Length 210;

Best Local Similarity 100.0%; Pred. No. 57;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 AELEQEV 26
|
|
|
|
|
|
|
Db 169 AELEQEV 175

Search completed: July 26, 2005, 12:28:17
Job time : 102.229 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 26, 2005, 12:00:54 ; Search time 19.7083 Seconds
(without alignments)
209.927 Million cell updates/sec

Title: US-10-624-218-4
Perfect score: 43
Sequence: 1 LEIPAAFLQRNTALRTEVA.....QEVORLENSQYETRYGPL 43

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 283416 seqs, 96216763 residues

Word size : 0
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : PIR_79:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	11	25.6	261	2 A40579	trans-activating t
2	11	25.6	293	2 S50109	vitellogenin gene-
3	11	25.6	303	2 G02360	thyrotroph embryon
4	11	25.6	313	2 A41524	transcription fact
5	7	16.3	74	2 B82597	hypothetical prote
6	7	16.3	216	2 G82421	conserved hypothet
7	7	16.3	220	2 D75589	lexA repressor - D
8	7	16.3	234	2 A82814	thiamin-phosphate
9	7	16.3	315	2 A84565	hypothetical prote
10	7	16.3	601	2 T37872	hypothetical serin
11	7	16.3	636	2 T35182	probable ABC-type
12	7	16.3	719	2 T35189	probable ATP-depen
13	7	16.3	1143	2 T22952	hypothetical prote
14	7	16.3	1186	2 T19050	hypothetical prote
15	7	16.3	2712	2 T05113	hypothetical prote
16	6	14.0	63	2 E41608	hypothetical prote
17	6	14.0	65	2 C90584	50S ribosomal prot
18	6	14.0	121	2 H71351	probable ribosomal
19	6	14.0	129	2 AF1177	transcription regu
20	6	14.0	159	2 D90485	hypothetical prote
21	6	14.0	173	2 S34767	neuropeptides prec
22	6	14.0	176	2 S52914	virion protein j13
23	6	14.0	183	2 S52904	virion protein j13
24	6	14.0	185	2 S52903	virion protein j13
25	6	14.0	189	2 PC1171	histo-blood group
26	6	14.0	189	2 PC1168	histo-blood group
27	6	14.0	189	2 PC1166	histo-blood group
28	6	14.0	189	2 PC1170	histo-blood group
29	6	14.0	189	2 PC1169	histo-blood group

30	6	14.0	189	2 PC1165	glycoprotein-fucos
31	6	14.0	189	2 PC1173	histo-blood group
32	6	14.0	189	2 PC1172	histo-blood group
33	6	14.0	193	2 F72311	lema protein - The
34	6	14.0	193	2 H91249	hypothetical prote
35	6	14.0	195	2 E75424	maf protein - Dein
36	6	14.0	197	1 K1PGGU	guanylate kinase (
37	6	14.0	197	2 S68864	guanylate kinase (
38	6	14.0	198	2 S39447	hypothetical prote
39	6	14.0	205	2 T19851	NADH oxidase homol
40	6	14.0	210	2 AH1087	hypothetical prote
41	6	14.0	228	2 D69317	probable GntR-fam1
42	6	14.0	230	2 H95265	probable p-aminobe
43	6	14.0	232	2 T10008	probable paba prot
44	6	14.0	232	2 C70699	hypothetical prote
45	6	14.0	234	2 T31753	hypothetical prote

ALIGNMENTS

RESULT 1

A40579
trans-activating transcription regulator TEF - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 05-Jun-1992 #sequence_revision 05-Jun-1992 #text_change 09-Jul-2004
C:Accession: A40579
R:Drolet, D.W.; Scully, K.M.; Simmons, D.M.; Wegner, M.; Chu, K.; Swanson, L.W.; Rosenf
Genes Dev. 5, 1739-1753, 1991
A>Title: TEF, a transcription factor expressed specifically in the anterior pituitary d
A:Reference number: A40579; MUID:92009166; PMID:1916262
A:Accession: A40579
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-261 <DRO>
A:Cross-references: UNIPROT:P41224; GB:S58745; NID:G237084; PIDN:AAB20032.1; PID:G23708-
C:Keywords: DNA binding; transcription regulation

Query Match 25.6%; Score 11; DB 2; Length 261;
Best Local Similarity 100.0%; Pred. No. 0.00069;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	12	NTALRTEVAEL	22
DB	230	NTALRTEVAEL	240

RESULT 2

S50109
vitellogenin gene-binding protein VBP, beta/beta isoform - chicken
C:Species: Gallus gallus (chicken)
C>Date: 14-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 09-Jul-2004
C:Accession: S50109; S50111
R:Burch, J.B.E.; Davis, D.L.
Nucleic Acids Res. 22, 4733-4741, 1994
A>Title: Alternative promoter usage and splicing options result in the differential exp
ors.
A:Reference number: S50109; MUID:95075656; PMID:7984425
A:Accession: S50109
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-233 <BUR>
A:Cross-references: UNIPROT:Q92172; EMBL:U09221; NID:G483937; PIDN:AAA82156.1; PID:G483
A:Accession: S50111
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-34 <BU2>
A:Cross-references: EMBL:U09223; PIDN:AAA82158.1; PID:G483824
C:Keywords: transcription factor

Query Match 25.6%; Score 11; DB 2; Length 293;
Best Local Similarity 100.0%; Pred. No. 0.00076;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 NTALRTEVAEL 22
| | | | | | | | | |
Db 256 NTALRTEVAEL 266

RESULT 3
G02360
thyrotroph embryonic factor - human
C:Species: Homo sapiens (man)
C:Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 09-Jul-2004
C:Accession: G02360; B55558
R:Hunger, S.P.
submitted to the EMBL Data Library, December 1995
A:Reference number: H01113
A:Accession: G02360
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-303 <HUN>
A:Cross-references: UNIPROT:Q8IU94; EMBL:U44059; NID:g1399385; PIDN:AA806497.1; PID:g139
R:Khatib, Z.A.; Inaba, T.; Valentine, M.; Look, A.T.
Genomic 23, 344-351, 1994
A:Title: Chromosomal localization and cDNA cloning of the human DBP and TEF genes.
A:Reference number: A55558; MUID:95137580; PMID:7835883
A:Accession: B55558
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 43-53, 'E', 55-303 <KHA>
A:Cross-references: GB:U06935; NID:g606796; PIDN:AAA81373.1; PID:g606797
C:Genetics:
A:Gene: GDB:TEF
A:Cross-references: GDB:359741; OMIM:188595
A:Map position: 22q13-22q13

Query Match 25.6%; Score 11; DB 2; Length 303;
Best Local Similarity 100.0%; Pred. No. 0.00079;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 NTALRTEVAEL 22
| | | | | | | | | |
Db 272 NTALRTEVAEL 282

RESULT 4
A41524
transcription factor, vitellogenin promoter-binding - chicken
C:Species: Gallus gallus (chicken)
C:Date: 28-May-1992 #sequence_revision 28-May-1992 #text_change 09-Jul-2004
C:Accession: A41524; S50110
R:Iyer, S.V.; Davis, D.L.; Seal, S.N.; Burch, J.B.E.
Mol. Cell. Biol. 11, 4863-4875, 1991
A:Title: Chicken vitellogenin gene-binding protein, a leucine zipper transcription facto
A:Reference number: A41524; MUID:92017760; PMID:1922023
A:Accession: A41524
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-313 <IYE>
A:Cross-references: UNIPROT:Q92057
R:Burch, J.B.E.; Davis, D.L.
Nucleic Acids Res. 22, 4733-4741, 1994
A:Title: Alternative promoter usage and splicing options result in the differential exp
ors.
A:Reference number: S50109; MUID:95075656; PMID:7984425
A:Accession: S50110
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-60 <BUR>
A:Cross-references: EMBL:U09222; NID:g483821; PIDN:AAA82157.1; PID:g483822

Query Match 25.6%; Score 11; DB 2; Length 313;
Best Local Similarity 100.0%; Pred. No. 0.00081;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 NTALRTEVAEL 22
| | | | | | | | | |
Db 282 NTALRTEVAEL 292

RESULT 5
B82597
hypothetical protein XF2130 [imported] - Xylella fastidiosa (strain 9a5c)
C:Species: Xylella fastidiosa
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C:Accession: B82597
R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; MUID:20365717; PMID:10910347
A:Note: for a complete list of authors see reference number A59328 below
A:Accession: B82597
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-74 <SIM>
A:Cross-references: UNIPROT:Q9PBL4; GB:AE004027; GB:AE003849; NID:g9107249; PIDN:AAF8492
A:Experimental source: strain 9a5c
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Canargo, L.E.A.; Carraro, D.M.; Carrer, H
as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigr
Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E.
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: XF2130

Query Match 16.3%; Score 7; DB 2; Length 74;
Best Local Similarity 100.0%; Pred. No. 3.1;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 RTEVAEL 22
| | | | | | | | | |
Db 17 RTEVAEL 23

RESULT 6
G82421
conserved hypothetical protein VCA0740 [imported] - Vibrio cholerae (strain N16961 serogru
C:Species: Vibrio cholerae
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C:Accession: G82421
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P
1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; MUID:20406833; PMID:10952301
A:Accession: G82421
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-216 <HEI>
A:Cross-references: UNIPROT:Q9KLK3; GB:AE004403; GB:AE003853; NID:g9658159; PIDN:AAF9663
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VCA0740
A:Map position: 2
C:Superfamily: Escherichia coli yceH protein

Query Match 16.3%; Score 7; DB 2; Length 216;
Best Local Similarity 100.0%; Pred. No. 8.2;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 AELEQEV 26
|||||

Db 191 AELEQEV 197
|||||

RESULT 7

D75589

lexA repressor - Deinococcus radiodurans (strain R1)

C:Species: Deinococcus radiodurans

C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000

C:Accession: D75589

R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.

Science 286, 1571-1577, 1999

A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.

A:Reference number: A75250; MUID:20036896; PMID:10567266

A:Accession: D75589

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-220 <WHI>

A:Cross-references: GB:AE001863; GB:AE001825; NID:G6460670; PIDN:AAF12438.1; PID:G646073

A:Experimental source: strain R1.

C:Genetics:

A:Gene: DRA0344

A:Map position: 2

C:Superfamily: lexA repressor

Query Match 16.3%; Score 7; DB 2; Length 220;
Best Local Similarity 100.0%; Pred. No. 8.3;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 37 ETRYGPL 43
|||||

Db 66 ETRYGPL 72
|||||

RESULT 8

A82814

thiamin-phosphate pyrophosphorylase XF0378 [imported] - Xylella fastidiosa (strain 9a5c)

C:Species: Xylella fastidiosa

C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004

C:Accession: A82814

R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequencing

Nature 406, 151-157, 2000

A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.

A:Reference number: A82515; MUID:20365717; PMID:10910347

A:Note: for a complete list of authors see reference number A59328 below

A:Accession: A82814

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-234 <SIM>

A:Cross-references: UNIPROT:O9PGC4; GB:AE003889; GB:AE003849; NID:g9105205; PIDN:AAF8318

A:Experimental source: strain 9a5c

R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H ab-Neco, E.; Docena, C.; El-Dorfy, H.; Facincani, A.P.; Ferreira, A.J.S.

submitted to GenBank, June 2000

A:Authors: Ferreira, V.C.A.; Perro, J.A.; Fraga, J.S.; Franco, M.C.; Franco, M.C.; Frohm J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigh Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.; F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir M.; Tuhako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z A:Reference number: A59328

A:Contents: annotation

C:Genetics:

A:Gene: XF0378

Query Match 16.3%; Score 7; DB 2; Length 234;
Best Local Similarity 100.0%; Pred. No. 8.8;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 12 NTALRTE 18
|||||

Db 72 NTALRTE 78
|||||

RESULT 9

A84565

hypothetical protein At2g18500 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004

C:Accession: A84565

R:Lin, X.; Hui, K.; Rounsley, S.D.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: A84420; MUID:20083487; PMID:10617197

A:Accession: A84565

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-315 <STO>

A:Cross-references: UNIPROT:Q9ZU65; GB:AE002093; NID:G4218008; PIDN:RAD12216.1; GSPDB:S A:Genetics:

A:Gene: At2g18500

A:Map position: 2

Query Match 16.3%; Score 7; DB 2; Length 315;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 EVALEEQ 24
|||||

Db 256 EVALEEQ 262
|||||

RESULT 10

T37872

hypothetical serine rich protein - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe

C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004

C:Accession: T37872

R:Skellton, J.; Churcher, C.M.; Barrall, B.G.; Rajandream, M.A.; Wood, V. submitted to the EMBL Data Library, August 1997

A:Reference number: Z21751

A:Accession: T37872

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-601 <SKE>

A:Cross-references: UNIPROT:O13803; EMBL:Z98597; PIDN:CAB11215.1; GSPDB:GN00066; SPDB:S: A:Experimental source: strain 972h-; cosmid c17H9

C:Genetics:

A:Gene: SPDB:SPAC17H9.06c

A:Map position: 1

A:Introns: 335/1; 500/3; 575/3

C:Superfamily: Schizosaccharomyces pombe hypothetical serine-rich protein SPAC17H9.06c

Query Match 16.3%; Score 7; DB 2; Length 601;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 ELEQEQV 27
|||||

Db 296 ELEQEQV 302
|||||

RESULT 11

T35182

probable ABC-type transport protein ramA - Streptomyces coelicolor

C:Species: Streptomyces coelicolor

C;Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 16-Aug-2004

C;Accession: T35182

R;Seeger, K.J.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, July 1998

A;Reference number: Z21570

A;Accession: T35182

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-636 <SEE>

A;Cross-references: UNIPROT:O88039; EMBL:AL031107; PIDN:CAA19962.1; GSPDB:GN00070; SCORE

A;Experimental source: strain A3(2)

C;Genetics:

A;Gene: rnaA; SCOREDB:SCSA7.33

C;Superfamily: ATP-binding cassette homology

Query Match 16.3%; Score 7; DB 2; Length 636;

Best Local Similarity 100.0%; Pred. No. 22;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 14 ALRTEVA 20

|||||

Db 424 ALRTEVA 430

RESULT 12

T35189

Probable ATP-dependent DNA helicase - Streptomyces coelicolor

C;Species: Streptomyces coelicolor

C;Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 09-Jul-2004

C;Accession: T35189

R;Seeger, K.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, April 1998

A;Reference number: Z21571

A;Accession: T35189

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-719 <SEE>

A;Cross-references: UNIPROT:O69992; EMBL:AL022374; PIDN:CAA18513.1; GSPDB:GN00070; SCORE

A;Experimental source: strain A3(2)

C;Genetics:

A;Gene: SCOREDB:SCSB8.05

Query Match

Best Local Similarity 16.3%; Score 7; DB 2; Length 719;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 AAFLRQR 11

|||||

Db 266 AAFLRQR 272

RESULT 13

T22952

hypothetical protein F58H1.1 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C;Accession: T22952

R;Harris, B.

submitted to the EMBL Data Library, July 1996

A;Reference number: Z19642

A;Accession: T22952

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-1143 <WIL>

A;Cross-references: UNIPROT:Q21010; EMBL:Z75954; PIDN:CAB00104.1; GSPDB:GN00023; CESP:F5

A;Experimental source: clone F58H1

C;Genetics:

A;Gene: CESP:F58H1.1

A;Map position: 5

A;Introns: 39/1; 142/3; 294/1; 328/2; 443/2; 517/1; 643/2; 702/3; 800/1; 1050/2

Query Match

Best Local Similarity 16.3%; Score 7; DB 2; Length 1143;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 22 LEQEVQR 28

|||||

Db 62 LEQEVQR 68

RESULT 14

T19050

hypothetical protein C07E3.3 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C;Accession: T19050

R;Matthews, P.

submitted to the EMBL Data Library, June 1995

A;Reference number: Z19066

A;Accession: T19050

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-1186 <WIL>

A;Cross-references: UNIPROT:Q17786; EMBL:Z49908; PIDN:CAA90094.1; GSPDB:GN00020; CESP:CO

A;Experimental source: clone C07E3

C;Genetics:

A;Gene: CESP:C07E3.3

A;Map position: 2

A;Introns: 36/3; 79/3; 1128/3; 1160/3

Query Match

Best Local Similarity 16.3%; Score 7; DB 2; Length 1186;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 22 LEQEVQR 28

|||||

Db 575 LEQEVQR 581

RESULT 15

T05113

hypothetical protein F28M20.240 - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004

C;Accession: T05113

R;Bavan, M.; Rieger, M.; Mueller-Auer, S.; Zipp, M.; Schaefer, M.; Hoheisel, J.; Mewes, J.

submitted to the Protein Sequence Database, November 1998

A;Reference number: Z15398

A;Accession: T05113

A;Molecule type: DNA

A;Residues: 1-2712 <BEV>

A;Cross-references: UNIPROT:Q9SB74; EMBL:AL031004

A;Experimental source: cultivar Columbia; BAC clone F28M20

C;Genetics:

A;Map position: 4

A;Introns: 17/3; 240/1; 1950/1; 2118/3; 2381/1; 2599/3; 2645/1; 2679/2

Query Match

Best Local Similarity 16.3%; Score 7; DB 2; Length 2712;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 ELEQEVQ 27

|||||

Db 2405 ELEQEVQ 2411

Search completed: July 26, 2005, 12:20:51

Job time : 20.7083 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 26, 2005, 11:59:44 ; Search time 97.3472 Seconds
(without alignments)
226.194 Million cell updates/sec

Title: US-10-624-218-4

Perfect score: 43

Sequence: 1 LEIRAAFLQRNTALTEVA.....QEVORLENSQVETRYGPL 43

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 1612378 seqs, 512079187 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : UniProt_03:*

1: uniprot_sprot:*

2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	11	25.6	140	2	Q723J7 homo sapien
2	11	25.6	271	1	TEF RAT
3	11	25.6	273	2	Q726X6 homo sapien
4	11	25.6	293	2	Q92172 gallus gall
5	11	25.6	301	1	TEF_MOUSE
6	11	25.6	303	1	TEF_HUMAN
7	9	20.9	221	1	TEF PHOSU
8	9	20.9	481	1	PUS3_HUMAN
9	8	18.6	429	2	Q7PRV5
10	7	16.3	74	2	Q9PBL4
11	7	16.3	92	2	Q7N2A2
12	7	16.3	104	2	Q83B11
13	7	16.3	124	2	Q89HV3
14	7	16.3	201	2	Q7NCR4
15	7	16.3	210	1	LEXA DEIRA
16	7	16.3	216	2	Q9KLK3
17	7	16.3	234	1	THIE_XYLFA
18	7	16.3	234	1	THIE_XYLFT
19	7	16.3	237	2	Q7PL35
20	7	16.3	252	2	Q7PPR9
21	7	16.3	315	2	Q9ZU65
22	7	16.3	342	2	Q67UN7
23	7	16.3	350	2	Q6N063
24	7	16.3	351	2	Q8S626
25	7	16.3	360	2	Q8KNB8
26	7	16.3	360	2	Q8G3C2
27	7	16.3	386	2	Q7PK66
28	7	16.3	393	2	Q9HK97
29	7	16.3	524	2	Q9LCB5
30	7	16.3	556	2	Q8T5S9
31	7	16.3	573	1	SUOX_DROME

32	7	16.3	593	2	Q9AS76	Q9as76 oryza sativ
33	7	16.3	601	1	YE06_SCHPO	O13803 schizosacch
34	7	16.3	615	2	Q9X618	Q9x618 streptomyce
35	7	16.3	636	2	Q53818	Q53818 streptomyce
36	7	16.3	636	2	O88039	O88039 streptomyce
37	7	16.3	719	2	O69992	O69992 streptomyce
38	7	16.3	725	2	Q8CB87	Q8cb87 mus musculu
39	7	16.3	752	2	Q9H0K0	Q9h0k0 homo sapien
40	7	16.3	847	2	Q6AFS4	Q6afs4 leifsonia x
41	7	16.3	882	2	Q8DLWS	Q8dlws synecococc
42	7	16.3	892	2	Q7VEG5	Q7veg5 prochloroco
43	7	16.3	895	2	Q8DM06	Q8dm06 synecococc
44	7	16.3	915	2	Q869R0	Q869r0 dictyosteli
45	7	16.3	942	2	Q658Z8	Q658z8 homo sapien

ALIGNMENTS

RESULT 1

Q723J7 PRELIMINARY; PRT; 140 AA.
 AC Q723J7; 01-OCT-2003 (Tremblrel. 25, Created)
 DT 01-OCT-2003 (Tremblrel. 25, Last sequence update)
 DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
 DE Hypothetical protein DKFZp686D1282 (Fragment).
 GN Name=DKFZp686D1282;
 OS Homo sapiens (Human);
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Human retina;
 RA Wambutt R., Haubner D., Mewes H.W., Weil B., Amid C., Osanger A.,
 Fobo G., Han M., Wiemann S.;
 RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: Belongs to the bZIP family.
 DR EMBL; BX537848; CAD97856.1; -;
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0003677; F:DNA binding; IEA.
 DR GO; GO:0006355; F:regulation of transcription, DNA-dependent; IEA.
 DR InterPro; IPR004827; TF_bZIP.
 DR SMART; SM00338; BRL2; 1.
 DR PROSITE; PS50217; BZIP; 1.
 KW DNA-binding; Hypothetical protein; Nuclear protein.
 FT NON_TER 1
 SQ SEQUENCE 140 AA; 16165 MW; EDD5A96CF193E0BC CRC64;

Query Match 25.6%; Score 11; DB 2; Length 140;
 Best Local Similarity 100.0%; Pred. No. 0.0049;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 NTALRTEVAEL 22
 |||||
 DB 109 NTALRTEVAEL 119

RESULT 2

TEF_RAT STANDARD; PRT; 271 AA.
 AC P41224;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 25-OCT-2004 (Rel. 45, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Thyrotroph embryonic factor (Fragment).
 GN Name=TEF;
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]

SEQUENCE FROM N.A., INTERACTION WITH DBP, AND MUTAGENESIS.
 TISSUE=Pituitary;
 MEDLINE=92009166; PubMed=1916262;
 RA Drolet D.W., Scully K.M., Simmons D.M., Wegner M., Chu K.,
 RA Swanson L.W., Rosenfeld M.G.;
 RT "TEF, a transcription factor expressed specifically in the anterior
 RT pituitary during embryogenesis, defines a new class of leucine zipper
 RT proteins.";
 RL Genes Dev. 5:1739-1753(1991).
 CC -1- FUNCTION: Transcription factor that binds to and transactivates
 CC the TSHB promoter. Binds to a minimal DNA-binding sequence 5'-
 CC [TC] [AG] [AG] [TTC] [AG]-3'.
 CC -1- SUBUNIT: Binds DNA as a homodimer or a heterodimer. Can form a
 CC heterodimer with DBP.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- TISSUE SPECIFICITY: Expressed exclusively in the rostral portion
 CC of the anterior pituitary during embryogenesis. Found in several
 CC tissues in juvenile and adult rats.
 CC -1- DEVELOPMENTAL STAGE: Expressed up to embryonic day 14 and
 CC specifically in the anterior pituitary during embryogenesis.
 CC -1- INDUCTION: Accumulates according to a robust circadian rhythm (By
 CC similarity).
 CC -1- SIMILARITY: Belongs to the bZIP family. PAR subfamily.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC -----
 DR EMBL; S58745; AAB20032.1; ALT_INIT.
 DR PIR; A40579; A40579.
 DR TRANSFAC; T01072; -.
 DR RGD; 3841; Tef.
 DR InterPro; IPR004827; TF_bZIP.
 DR Pfam; PF00170; bZIP; 1.
 DR SMART; SM00336; BRLZ; 1.
 DR PROSITE; PS00217; bZIP; 1.
 DR PROSITE; PS0036; bZIP_BASIC; FALSE NEG.
 KW Activator; Biological rhythms; DNA-binding; Nuclear protein;
 KW Transcription regulation.
 FT NON_TER 1
 FT DOMAIN 134 185 Pro-rich (Proline/Acidic region (PAR)).
 FT DNA_BIND 216 225 Basic motif.
 FT DOMAIN 236 250 K->A: 30-fold decrease in affinity for
 FT MUTAGEN 188 188 prolactin recognition element.
 FT MUTAGEN 189 189 prolactin recognition element.
 FT MUTAGEN 191 191 K->A: 30-fold decrease in affinity for
 FT MUTAGEN 192 192 prolactin recognition element.
 FT MUTAGEN 192 192 K->A: 30-fold decrease in affinity for
 FT MUTAGEN 236 236 prolactin recognition element.
 FT MUTAGEN 243 243 L->V: Diminishes DNA-binding.
 FT SEQUENCE 271 AA; 30330 MW; 1ASCFA2396C1188B9 CRC64;
 Query Match 25.6%; Score 11; DB 1; Length 271;
 Best Local Similarity 100.0%; Pred. No. 0.0089;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 12 NTALRTEVAEL 22
 DB 240 NTALRTEVAEL 250
 RESULT 3
 Q726X6 PRELIMINARY; PRT; 273 AA.
 AC Q726X6;
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)

DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE DJ95NI.5.2 (Tnyrotrophic embryonic factor (Ortholog of chicken
 DE vitellogenin gene-binding protein VBP beta/beta variant) (Variant
 DE 2)).
 GN Name=TEF;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Smith M.;
 RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: Belongs to the bZIP family.
 CC EMBL; AL035659; CAB62497.1; -
 CC GO; GO:0005634; C:nucleus; IEA.
 CC GO; GO:0003677; F:DNA binding; IEA.
 CC GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 CC InterPro; IPR004827; TF_bZIP.
 CC SMART; SM00338; BRLZ; 1.
 CC PROSITE; PS00217; bZIP; 1.
 KW DNA-binding; Nuclear protein.
 SQ SEQUENCE 273 AA; 30634 MW; B1A956839C8C4AC4 CRC64;
 Query Match 25.6%; Score 11; DB 2; Length 273;
 Best Local Similarity 100.0%; Pred. No. 0.009;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 12 NTALRTEVAEL 22
 DB 242 NTALRTEVAEL 252
 RESULT 4
 Q92172 PRELIMINARY; PRT; 293 AA.
 AC Q92172;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Vitellogenin binding protein (VBP), beta/beta isoform.
 GN Name=VBP;
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN=White Leghorn;
 RC MEDLINE=95075656; PubMed=7984425;
 RX Burch J.B., Davis D.L.;
 RA "Alternative promoter usage and splicing options result in the
 RT differential expression of mRNAs encoding four isoforms of chicken
 RT VBP, a member of the PAR subfamily of bZIP transcription factors.";
 RL Nucleic Acids Res. 22:4733-4741(1994).
 CC -1- SIMILARITY: Belongs to the bZIP family.
 CC EMBL; U09221; AAA82156.1; -
 CC PIR; S50109; S50109.
 CC GO; GO:0005634; C:nucleus; IEA.
 CC GO; GO:0003677; F:DNA binding; IEA.
 CC GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 CC SMART; SM00338; BRLZ; 1.
 CC PROSITE; PS00217; bZIP; 1.
 KW DNA-binding; Nuclear protein.
 SQ SEQUENCE 293 AA; 32736 MW; C5803D510CE803D6 CRC64;
 Query Match 25.6%; Score 11; DB 2; Length 293;
 Best Local Similarity 100.0%; Pred. No. 0.0096;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 12 NTALRTEVAEL 22

|||||||
256 NTALRTEVAEL 266

Db

RESULT 5

TEF MOUSE

ID - TEF MOUSE STANDARD; PRT; 301 AA.

AC Q9JUC6; Q6QHT6; Q8C610; Q8VD02;

DT 25-OCT-2004 (Rel. 45, Created)

DT 25-OCT-2004 (Rel. 45, Last sequence update)

DT 25-OCT-2004 (Rel. 45, Last annotation update)

DE Tyrotroph embryonic factor.

GN Name=TEF;

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN

RN SEQUENCE FROM N.A. (ISOFORM ALPHA).

RP STRAIN=BALE/C; TISSUE=Brain;

RP MEDLINE=20261582; PubMed=10799536; DOI=10.1074/jbc.275.19.14524;

RA Krueger D.A., Warner E.A., Dowd D.R.;

RT "Involvement of thyrotroph embryonic factor in calcium-mediated

RT regulation of gene expression.";

RL J. Biol. Chem. 275:14524-14531 (2000).

RN

RN [2]

RP SEQUENCE FROM N.A. (ISOFORMS ALPHA AND BETA), AND ALTERNATIVE PROMOTER

RP USAGE.

RX PubMed=14702338; DOI=10.1074/jbc.M313822200;

RA Zhou J., Hoggatt A.M., Herring B.P.;

RT "Activation of the smooth muscle-specific telokin gene by thyrotroph

RT embryonic factor (TEF).";

RL J. Biol. Chem. 279:15929-15937 (2004).

RN

RN [3]

RP SEQUENCE FROM N.A. (ISOFORMS ALPHA AND 2).

RP STRAIN=FVB/N; TISSUE=Breast tumor, and Kidney;

RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Shenmen C.M., Schuler G.D.,

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Brownstein M.J., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,

RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huliy S.W.,

RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,

RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,

RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length human

RT and mouse cDNA sequences.";

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).

RN

RN [4]

RP SEQUENCE OF 97-196 FROM N.A.

RX STRAIN=C57BL/6J; TISSUE=Kidney;

RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;

RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,

RA Nikaio I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,

RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,

RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,

RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,

RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,

RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,

RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,

RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,

RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,

RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,

RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,

RA Nagashima T., Nunata K., Okido T., Pavan W.J., Pertea G., Pesole G.,

RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramchandran S.,

RA Ravasi T., Reed J.C., Reid J., Reid J., Ring B.Z., Ringwald M.,

RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,

RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,

RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,

RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,

RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,

RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,

RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,

RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,

RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,

RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,

RA Birney E., Hayashizaki Y.;

RT "Analysis of the mouse transcriptome based on functional annotation of

RT 60,770 full-length cDNAs.";

RL Nature 420:563-573 (2002).

RN

RN [5]

RP CIRCADIAN INDUCTION.

RX PubMed=8617210;

RA Fonjallaz P., Ossipow V., Wanner G., Schibler U.;

RT "The two PAR leucine zipper proteins, TEF and DBP, display similar

RT circadian and tissue-specific expression, but have different target

RT promoter preferences.";

RL EMBO J. 15:351-362 (1996).

RN

RN [6]

RP IMPLICATION IN EPILEPSY.

RX PubMed=15175240; DOI=10.1101/gad.301404;

RA Gachon F., Fonjallaz P., Damiola F., Gos P., Kodama T., Zakany J.,

RA Duboule D., Petit B., Tafti M., Schibler U.;

RT "The loss of circadian PAR bzip transcription factors results in

RT epilepsy.";

RL Genes Dev. 18:1397-1412 (2004).

CC -!- FUNCTION: Transcription factor that binds to and transactivates

CC the TSHB promoter. Binds to a minimal DNA-binding sequence 5'-

CC [TC][AG][AG][TTA][TC][AG]-3' (By similarity). Also activates the

CC telokin promoter in smooth muscle-specific and calcium-dependent

CC manner.

CC -!- SUBUNIT: Binds DNA as a homodimer or a heterodimer. Can form a

CC heterodimer with DBP (By similarity).

CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).

CC -!- ALTERNATIVE PRODUCTS:

CC Event-Alternative promoter;

CC Comment=3 isoforms, Alpha (shown here), Beta and 2, are produced

CC by use of alternative promoters;

CC -!- TISSUE SPECIFICITY: Isoform Alpha and isoform Beta are expressed

CC at high levels in lung, bladder, kidney, gut and brain.

CC -!- INDUCTION: Accumulates according to a robust circadian rhythm in

CC liver and kidney. In liver nuclei, the amplitude of daily

CC oscillation has been estimated to be 9-fold. Expressed at nearly

CC constant level in the brain.

CC -!- MISCELLANEOUS: Mice deficient for all three PAR bzip proteins

CC (DBP, HLF and TEF) display a dramatically shortened life span and

CC are highly susceptible to generalized spontaneous and audiogenic

CC epilepsies (due for example to the noise of a vacuum cleaner) that

CC are frequently lethal. The down-regulation of pyridoxal kinase

CC (Pdxk) expression in these mice may participate in this seizure

CC phenotype.

CC -!- SIMILARITY: Belongs to the bzip family. PAR subfamily.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>).

CC or send an email to license@isb-sib.ch.

CC

CC EMBL; AF194420; AAF70830.1; -

CC EMBL; AY540631; AAS45599.1; -

CC EMBL; AY540632; AAS45600.1; -

CC EMBL; BC017689; AAH17689.1; -

CC EMBL; BC036982; AAH36982.1; -

```
DR EMBL; AK075601; BAC35849.1; -.
DR MGD; MGI:98663; Tef.
DR GO; GO:0005634; C:nucleus; IC.
DR GO; GO:0003690; C:double-stranded DNA binding; IDA.
DR GO; GO:0005515; F:protein binding; IPI.
DR GO; GO:0016563; F:transcriptional activator activity; IDA.
DR GO; GO:0045944; P:positive regulation of transcription from P. . .; IDA.
DR InterPro; IPR004827; TF_bZIP.
DR Pfam; PF00170; bZIP; 1.
DR SMART; SM00338; BRLZ; 1.
DR PROSITE; PS00217; bZIP; 1.
DR PROSITE; PS00036; bZIP_BASIC; FALSE NEG.
KW Activator; Alternative promoter usage; Biological rhythms;
KW DNA-binding; Nuclear protein; Transcription regulation.
FT DOMAIN 164 215 Pro-rich (Proline/Acidic region (PAR)).
FT DNA_BIND 237 255 Basic motif.
FT DOMAIN 266 280 Leucine-zipper.
FT VARSPLIC 1 50 MSDAGGKKPVEFQAGPGRAAGRLSGSFLVLKML
FT FT ENPPRETRL -> MSSCSQIGVAPADMPEVLKSLLEHSLP
FT FT WSEKKA (in isoform Beta).
FT FT /FTId=VSP_011245.
FT VARSPLIC 1 50 MSDAGGKKPVEFQAGPGRAAGRLSGSFLVLKML
FT FT ENPPRETRL -> MDMPVLKSLLEHSLPWSKKA (in
FT FT isoform 2).
FT FT /FTId=VSP_011246.
FT FT M -> V (in Ref. 4).
FT CONFLICT 97 97 P33145 MW; P33FFFC6D7E09144 CRC64;
SQ SEQUENCE 301 AA; 33145 MW; P33FFFC6D7E09144 CRC64;
Query Match 25.6%; Score 11; DB 1; Length 301;
Best Local Similarity 100.0%; Pred. No. 0.0098;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 12 NTALRTEVAEL 22
Db 270 NTALRTEVAEL 280
|||||
RESULT 6
TEF_HUMAN
ID -TEF_HUMAN STANDARD; PRT; 303 AA.
AC Q10587; Q15729; Q81U94; Q96TG4;
DT 01-OCT-1996 (Rel. 34, Created)
DT 25-OCT-2004 (Rel. 45, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Tyrocytoph embryonic factor.
GN Name=TEF;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=95137580; PubMed=7835883;
RA Khatib Z.A., Inaba T., Valentine M., Look A.T.;
RT "Chromosomal localization and cDNA cloning of the human DBP and TEF
genes."
RL Genomics 23:344-351(1994).
RN [2]
SEQUENCE FROM N.A.
RX MEDLINE=96219638; PubMed=8639829;
RA Hunger S.P., Li S., Fall M.Z., Naumovski L., Cleary M.L.;
RT "The proto-oncogene HLF and the related basic leucine zipper protein
TEF display highly similar DNA-binding and transcriptional regulatory
properties."
RL Blood 87:4607-4617(1996).
RN [3]
SEQUENCE FROM N.A.
RX COLLINS J.E., Wright C.L., Edwards C.A., Davis M.P., Grinham J.A.,
RA Cole C.G., Goward M.E., Aguado B., Mallia M., Mokrab Y., Huckle E.J.,
RA Beare D.M., Dunham I.;
RT "A genome annotation driven approach to cloning the human ORFeome."
RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.
RN [4]
```

```
RP SEQUENCE FROM N.A.
RA Halleck A., Ebert L., Moundinya M., Schick M., Eisenstein S.,
RA Neubert P., Kstrang K., Schatten R., Shen B., Henze S., Mar W.,
RA Korn B., Zuo D., Hu Y., LaBaer J.;
RT "Cloning of human full open reading frames in Gateway(TM) system entry
vector (pDONR201).";
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RX MEDLINE=20057165; PubMed=10591208; DOI=10.1038/990031;
RA Dunham I., Hunt A.R., Collins J.E., Bruskewich R., Beare D.M.,
RA Clamp M., Smink L.J., Ainscough R., Almeida J.P., Babbage A.K.,
RA Bagguley C., Bailey J., Barlow K.F., Bates K.N., Beasley O.P.,
RA Bird C.P., Blakey S.E., Bridgeman A.M., Buck D., Burgess J.,
RA Burrill W.D., Burton J., Carder C., Carter N.P., Chen Y., Clark G.,
RA Clegg S.M., Cobley V.E., Cole C.G., Collier R.E., Connor R.,
RA Conroy D., Corby N.R., Coville G.J., Cox A.V., Davis J., Dawson E.,
RA Dhani P.D., Dockree C., Dodsworth S.J., Durbin R.M., Ellington A.G.,
RA Evans K.L., Fey J.M., Fleming K., French L., Garner A.A.,
RA Gilbert J.G.R., Goward M.E., Grafham D.V., Griffiths M.N.D., Hall C.,
RA Hall R.E., Hall-Tanlyn G., Heathcote R.W., Ho S., Holmes S.,
RA Hunt S.E., Jones M.C., Kershaw J., Kimberley A.M., King A.,
RA Laird G.K., Langford C.F., Leversha M.A., Lloyd C., Lloyd D.M.,
RA Martyn I.D., Mashreghi-Mohammadi M., Matthews L.H., Mccann O.T.,
RA Mcclay J., McLaren S., McMurray A.A., Milne S.A., Mortimore B.J.,
RA Odell C.N., Pavitt R., Pearce A.V., Pearson D., Phillimore B.J.C.T.,
RA Phillips S.H., Plumb R.W., Ramsay H., Ramsey Y., Rogers L., Ross M.T.,
RA Scott C.E., Sehra H.K., Skuce C.D., Smalley S., Smith M.L., Swann R.M.,
RA Suderlund C., Spragon L., Stewart C.A., Sulston J.E., Swann R.M.,
RA Vaudin M., Wall M., Wallis J.M., Whiteley M.N., Willey D.L.,
RA Williams L., Williams S.A., Williamson H., Wilmer T.E., Wilming L.,
RA Wright C.L., Hubbard T., Bentley D.R., Beck S., Rogers J., Shimizu N.,
RA Minoeshima S., Kawasaki K., Sasaki T., Asakawa S., Kudoh J.,
RA Shintani A., Shibuya K., Yoshizaki Y., Aoki N., Mitsuyama S.,
RA Roe B.A., Chen F., Chu L., Crabtree J., Deschamps S., Do A., Do T.,
RA Dorman A., Fang F., Fu Y., Hu P., Hua A., Kenton S., Lai H., Lao H.I.,
RA Lewis J., Lewis S., Lin S.-P., Loh P., Malaj E., Nguyen T., Pan H.,
RA Phan S., Qi S., Qian Y., Ray L., Ren Q., Shaull S., Sloan D., Song L.,
RA Wang Q., Wang Y., Wang Z., White J., Willingham D., Wu H., Yao Z.,
RA Zhan M., Zhang G., Chisoe S., Murray J., Miller N., Minx P.,
RA Fulton R., Johnson D., Bemis G., Bentley D., Bradshaw H., Bourne S.,
RA Cordes M., Du Z., Fulton L., Goela D., Graves T., Hawkins J.,
RA Hinds K., Kemp K., Latreille P., Layman D., Ozerky P., Rohlfing T.,
RA Scheet P., Walker C., Wamsley A., Wohlmann P., Pepin K., Nelson J.,
RA Korf I., Bedell J.A., Hillier L.W., Mardis E., Waterston R.,
RA Wilson R., Emanuel B.S., Shaikh T., Kurahashi H., Saitta S.,
RA Budarf M.L., McDermid H.E., Johnson A., Wong A.C.C., Morrow B.E.,
RA Edelmann L., Kim U.J., Shizuya H., Simon M.I., Dumanski J.P.,
RA Peyrard M., Kedra D., Seroussi E., Franconi I., Tapia I., Bruder C.E.,
RA O'Brien K.P., Wilkinson P., Bodenteich A., Hartman K., Hu X.,
RA Khan A.S., Lane L., Tilahun Y., Wright H.;
RT "The DNA sequence of human chromosome 22."
RL Nature 402:489-495(1999).
RN [6]
RP SEQUENCE FROM N.A.
RX TISSUE=Brain;
MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan B., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.A., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Basak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Maman A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
```

RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RL and mouse cDNA sequences.";
 CC Proc. Natl. Acad. Sci. U.S.A. 99:16999-16903(2002).
 CC -!- FUNCTION: Transcription factor that binds to and transactivates
 CC the TSHB promoter. Binds to a minimal DNA-binding sequence 5'-
 CC [TC][AG][AG]TTA[TC][AG]-3'. (By similarity).
 CC -!- SUBUNIT: Binds DNA as a homodimer or a heterodimer. Can form a
 CC heterodimer with DBP.
 CC -!- SUBUNIT: Binds DNA as a homodimer or a heterodimer. Can form a
 CC heterodimer with DBP.
 CC -!- SUBCELLULAR LOCATION: Nuclear.
 CC -!- INDUCTION: Accumulates according to a robust circadian rhythm (By
 CC similarity).
 CC -!- SIMILARITY: Belongs to the bZIP family. PAR subfamily.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; U06935; AAA81373.1; ALT_INIT.
 CC EMBL; U44059; AAB06497.1; -.
 CC EMBL; CR456592; CAG30478.1; -.
 CC EMBL; CR541827; CAG46626.1; -.
 CC EMBL; AL035659; CAB62498.1; ALT_INIT.
 CC EMBL; BC039258; AAH39258.1; -.
 CC EMBL; BC042476; AAH42476.1; -.
 CC PIR; G02360; G02360.
 CC TRANSFAC; T04876; -.
 CC Genew; HGNC:11722; TEF.
 CC
 CC MIM; 188595; -.
 CC GO; GO:0003702; F:RNA polymerase II transcription factor acti. .; TAS.
 CC GO; GO:0006357; F:regulation of transcription from Pol II pro. .; TAS.
 CC InterPro; IPR004827; TF_bZIP.
 CC Pfam; PF00170; bZIP; 1.
 CC SMART; SM00338; BRLZ; 1.
 CC PROSITE; PS0217; bZIP; 1.
 CC PROSITE; PS00036; bZIP_BASIC; FALSE NEG.
 CC Activator; Biological rhythms; DNA-binding; Nuclear protein;
 CC Transcription regulation.
 CC DOMAIN 166 217 Pro-rich (Proline/Acidic region (PAR)).
 CC DNA BIND 239 257 Basic motif.
 CC DOMAIN 268 282 Leucine-zipper.
 CC CONFLICT 54 54 K -> E (in Ref. 1).
 CC SEQUENCE 303 AA; 33247 MW; 4A87B7BFA7248C6F CRC64;
 CC
 CC Query Match 25.6%; Score 11; DB 1; Length 303;
 CC Best Local Similarity 100.0%; Pred. No. 0.0099;
 CC Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC
 CC Qy 12 NTALRTEVAEL 22
 CC |||||
 CC Db 272 NTALRTEVAEL 282
 CC
 CC RESULT 7
 CC TEF PHOSU
 CC ID - TEF PHOSU STANDARD; PRT; 221 AA.
 CC AC P97516;
 CC DT 30-MAY-2000 (Rel. 39, Created)
 CC DT 30-MAY-2000 (Rel. 39, Last sequence update)
 CC DT 25-OCT-2004 (Rel. 45, Last annotation update)
 CC DE Thyrotroph embryonic factor (Fragment).
 CC GN Name=TEF;
 CC OS Phodopus sungorus (Striped hairy-footed hamster) (Djungarian hamster).
 CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
 CC OC Phodopus.
 CC OX NCBI_TaxID=10044;
 CC RN [1]
 CC SEQUENCE FROM N.A.

RA Bockmann J.;
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: Transcription factor that binds to and transactivates
 CC the TSHB promoter. Binds to a minimal DNA-binding sequence 5'-
 CC [TC][AG][AG]TTA[TC][AG]-3'. (By similarity).
 CC -!- SUBUNIT: Binds DNA as a homodimer or a heterodimer. Can form a
 CC heterodimer with DBP (By similarity).
 CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
 CC -!- INDUCTION: Accumulates according to a robust circadian rhythm.
 CC -!- SIMILARITY: Belongs to the bZIP family. PAR subfamily.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; Y11149; CAA72036.1; -.
 CC InterPro; IPR004827; TF_bZIP.
 CC SMART; SM00338; BRLZ; 1.
 CC PROSITE; PS0217; bZIP; 1.
 CC PROSITE; PS00036; bZIP_BASIC; FALSE NEG.
 CC Activator; Biological rhythms; DNA-binding; Nuclear protein;
 CC Transcription regulation.
 CC NON TER 1 1
 CC DOMAIN 106 157 Pro-rich (Proline/Acidic region (PAR)).
 CC DNA BIND 179 197 Basic motif.
 CC DOMAIN 208 >221 Leucine-zipper.
 CC NON TER 221 221
 CC SEQUENCE 221 AA; 24528 MW; 69645BCA042CFA8B CRC64;
 CC
 CC Query Match 20.9%; Score 9; DB 1; Length 221;
 CC Best Local Similarity 100.0%; Pred. No. 0.73;
 CC Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC
 CC Qy 12 NTALRTEVA 20
 CC |||||
 CC Db 212 NTALRTEVA 220
 CC
 CC RESULT 8
 CC PUS3 HUMAN
 CC ID - PUS3 HUMAN STANDARD; PRT; 481 AA.
 CC AC Q9BZE2; Q96D17; Q96J23; Q96NB4;
 CC DT 25-OCT-2004 (Rel. 45, Created)
 CC DT 25-OCT-2004 (Rel. 45, Last sequence update)
 CC DT 25-OCT-2004 (Rel. 45, Last annotation update)
 CC DE Pseudouridylylase synthase 3 (EC 4.2.1.70) (Pseudouridine synthase 3)
 CC DE (FKSG32 protein).
 CC GN Name=PUS3; Synonyms=FKSG32;
 CC OS Homo sapiens (Human).
 CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 CC OX NCBI_TaxID=9606;
 CC RN [1]
 CC SEQUENCE FROM N.A.
 CC Wang Y.-G., Gong L.;
 CC RA "Identification and characterization of FKSG32, a novel gene expressed
 CC in lung carcinoma.";
 CC RT Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 CC
 CC RL [2]
 CC SEQUENCE FROM N.A.
 CC RX PubMed=14702039; DOI=10.1038/ng1285;
 CC RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
 CC RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
 CC RA Sekine M., Ohtsuka M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
 CC RA Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,
 CC RA Nagahari K., Murakami K., Yasuda T., Watanabe T., Wagatsuma M.,
 CC RA Shiratori K., Sudo H., Hosoi T., Kaku Y., Kodaira H., Kondo H.,
 CC RA Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E.,
 CC RA Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M.,

RA Simpson A.J.G., Reinach P.C., Arruda P., Abreu F.A., Acencio M.,
RA Alvarenga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carriaro D.M., Carier H.,
RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,
RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferio J.A.,
RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,
RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
RA Lenos E.G.M., Lenos M.V.P., Lopes S.A., Lopes C.R., Machado J.A.,
RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA da Silva J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tshako M.H.,
RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
RT "The genome sequence of the plant pathogen *Xylella fastidiosa*."
RL Nature 406:151-159(2000).
DR EMBL; AE004027; AAF84929.1; --
DR PIR; B82597; B82597.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 74 AA; 8148 MW; 6310B64F6A230D1B CRC64;

Query Match 16.3%; Score 7; DB 2; Length 74;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 RTEVAEL 22
Db 17 RTEVAEL 23
|||||

RESULT 11
Q7N2A2 PRELIMINARY; PRT; 92 AA.
AC Q7N2A2;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Similar to unknown protein Yfcl of *Escherichia coli*.
GN OrderedLocusNames=plu3186;
OS Photorhabdus luminescens (subsp. laumondii).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Photorhabdus.
OX NCBI_TaxID=141679;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TT01;
RX MEDLINE=22957627; PubMed=14528314;
RA Duchaud E., Rusniok C., Frangeul L., Buchrieser C., Givaudan A.,
RA Taourit S., Bocs S., Boursaux-Eude C., Chandler M., Charles J.-P.,
RA Dassa E., Derosé R., Derzelle S., Freyssinet G., Gaudriault S.,
RA Medigue C., Lanois A., Powell K., Sigulier P., Vincent R., Wingate V.,
RA Zouine M., Glaser P., Boemare N., Danchin A., Kunst F.;
RT "The genome sequence of the entomopathogenic bacterium *Photorhabdus luminescens*."
RL Nat. Biotechnol. 21:1307-1313(2003).
RL EMBL; BX571869; CAE15560.1; --
DR PhotoList; plu3186; --
DR InterPro; IPR009098; Pepsin_inhib_3.
KW Complete proteome.
SQ SEQUENCE 92 AA; 10223 MW; 1CB8C6F30BB6BEC5 CRC64;

Query Match 16.3%; Score 7; DB 2; Length 92;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 VAELEQE 25
Db 39 VAELEQE 45
|||||

RESULT 12
Q83BI1 PRELIMINARY; PRT; 104 AA.
AC Q83BI1;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=CBU1527;
OS *Coxiella burnetii*.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Legionellales;
OC Coxiellaceae; *Coxiella*.
OX NCBI_TaxID=777;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Nine Mile phase I / RSA 493;
RX MEDLINE=22608657; PubMed=12704232; DOI=10.1073/pnas.0931379100;
RA Sehadrri R., Paulsen I.T., Eisen J.A., Read T.D., Nelson K.E.,
RA Nelson W.C., Ward N.L., Tettelin H., Davidson T.M., Beanan M.J.,
RA DeBoy R.T., Daugherty S.C., Brinkac L.M., Madupu R., Dodson R.J.,
RA Khouri H.M., Lee K.H., Carty H.A., Scanlan D., Heinzen R.A.,
RA Thompson H.A., Samuel J.E., Fraser C.M., Heidelberg J.F.;
RT "Complete genome sequence of the Q-fever pathogen, *Coxiella burnetii*."
RL Proc. Natl. Acad. Sci. U.S.A. 100:5455-5460(2003).
DR EMBL; AE016964; AA091024.1; --
DR TIGR; CBU1527; --
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 104 AA; 11703 MW; CD3AC647D80252FF CRC64;

Query Match 16.3%; Score 7; DB 2; Length 104;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 RTEVAEL 22
Db 70 RTEVAEL 76
|||||

RESULT 13
Q89HV3 PRELIMINARY; PRT; 124 AA.
AC Q89HV3;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Two-component hybrid sensor and regulator.
GN OrderedLocusNames=bll5886;
OS *Bradyrhizobium japonicum*.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; *Bradyrhizobium*.
OX NCBI_TaxID=375;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=USDA110;
RX MEDLINE=22484998; PubMed=12597275;
RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,
RA Sasamoto S., Watanabe A., Ideawa K., Iriguchi M., Kawashima K.,
RA Kohara M., Matsumoto M., Shimpo S., Tsuruoka H., Wada T., Yamada M.,
RA Tabata S.;
RA "Complete genomic sequence of nitrogen-fixing symbiotic bacterium *Bradyrhizobium japonicum* USDA110."
RT DNA Res. 9:189-197(2002).
RL EMBL; AP005956; BAC51151.1; --

```

DR GO: GO:0000155; F:Two-component sensor molecule activity; IEA.
DR GO: GO:000160; P:Two-component signal transduction system (p. . .; IEA.
DR InterPro; IPR008207; Hpt.
DR Pfam; PF01627; Hpt; 1.
DR SMART; SM00073; HPT; 1.
KW Complete proteome.
SQ SEQUENCE 124 AA; 13261 MW; 5C820E6EA39F8655 CRC64;

Query Match 16.3%; Score 7; DB 2; Length 124;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 AELEQEV 26
Db 36 AELEQEV 42

RESULT 14
Q7NCR4 PRELIMINARY; PRT; 201 AA.
AC Q7NCR4;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE G12914 protein.
GN OrderedLocusNames=g12914;
OS Gloeobacter violaceus.
OC Bacteria; Cyanobacteria; Chroococcales; Gloeobacter.
OX NCBI_TaxID=33072;
RN SEQUENCE FROM N.A.
RC STRAIN=PCC 7421;
RX MEDLINE=22977040; PubMed=14621292;
RA Nakamura Y., Kaneko T., Sato S., Mimuro M., Miyashita H., Tsuchiya T.,
RA Sasamoto S., Watanabe A., Kawashima K., Kishida Y., Kiyokawa C.,
RA Kohara M., Matsumoto M., Matsuno A., Nakazaki N., Shimpo S.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of Gloeobacter violaceus PCC 7421, a
RL DNA Res. 10:137-145(2003).
DR EMBL; AP006578; BAC90855.1; -.
KW Complete proteome.
SQ SEQUENCE 201 AA; 21538 MW; CEC9EF4A6AF82B64 CRC64;

Query Match 16.3%; Score 7; DB 2; Length 201;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 ALRTEVA 20
Db 19 ALRTEVA 25

RESULT 15
LEXA_DEIRA STANDARD; PRT; 210 AA.
AC Q32506;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE LexA repressor (EC 3.4.21.88).
GN Names=LexA; OrderedLocusNames=DRA0344;
OS Deinococcus radiodurans.
OC Bacteria; Deinococcus-Thermus; Deinococci; Deinococcales;
OC Deinococcaceae; Deinococcus.
OX NCBI_TaxID=1299;
RN SEQUENCE FROM N.A.
RC STRAIN=KD8301;
RA Narumi I., Kong X., Du Z., Cherchu K., Kitayama S., Watanabe H.;
RT "Cloning, sequencing and expression of the LexA-like gene of
RT Deinococcus radiodurans."
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.

```

```

RN SEQUENCE FROM N.A.
RP STRAIN=R1 / ATCC 13939 / DSM 20539 / NCIB 9279;
RX MEDLINE=20036896; PubMed=10567266; DOI=10.1126/science.286.5444.1571;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L.A., Utterback T.R., Zalewski C.,
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S.L., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RT "Genome sequence of the radioresistant bacterium Deinococcus
RT radiodurans R1."
RL Science 286:1571-1577(1999).
CC -!- FUNCTION: Represses a number of genes involved in the response to
CC DNA damage (SOS response), including recA and LexA. In the
CC presence of single-stranded DNA, recA interacts with LexA causing
CC an autocatalytic cleavage which disrupts the DNA-binding part of
CC LexA, leading to derepression of the SOS regulon and eventually
CC DNA repair (By similarity).
CC -!- CATALYTIC ACTIVITY: Hydrolysis of Ala-Gly bond in repressor
CC LexA.
CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- SIMILARITY: Belongs to the peptidase S24 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AB003475; BAA21376.1; -.
DR EMBL; AB001863; AAF12438.1; ALT_INIT.
DR HSSP; P03033; 1JHC.
DR MEROPS; S24.UPW; -.
DR TIGR; DRA0344; -.
DR HAMAP; MF 00015; -.
DR InterPro; IPR006198; Pept_S24_S26.
DR InterPro; IPR011056; Pept_S24_S26_C.
DR InterPro; IPR006197; Pept_S24_SOS.
DR InterPro; IPR009058; Wing_hix_DNA_bnd.
DR Pfam; PF00717; Peptidase_S24; 1.
DR PRINTS; PR00726; LEXASERPTASE.
DR Autocatalytic cleavage; Complete proteome; DNA damage;
DR DNA replication; DNA-binding; Hydrolase; Repressor; SOS response;
KW Transcription regulation.
KW DNA_BIND 25 44 H-T-H motif.
FT SITE 84 85 Cleavage (auto-) (By similarity).
FT ACT_SITE 120 120 Involved in auto-cleavage (By
FT ACT_SITE 159 159 Involved in auto-cleavage (By
FT similarity).
SQ SEQUENCE 210 AA; 22351 MW; 2DCEE8C87DD9331 CRC64;

Query Match 16.3%; Score 7; DB 1; Length 210;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 37 ETRYGFL 43
Db 56 ETRYGFL 62

Search completed: July 26, 2005, 12:19:41
Job time : 99.3472 secs

```


Query Match: 25.58% Indels: 0
DB: 6 Gaps: 0

US-10-624-218-4 (1-43) x AX305813 (1-446)

Qy 12 AsnThrAlaLeuArgThrGluValAlaGluLeu 22
|||||
Db 112 AACACAGCCTCGCGAGGAGTTGCCGAGCTT 144

RESULT 2
S58745
LOCUS
DEFINITION
thyrotroph embryonic factor-leucine zipper transcription factor
[rats, pituitary, mRNA, 817 nt].
S58745
S58745.1 GI:237084
Rattus sp.
ORGANISM
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
1 (bases 1 to 817)
Drolet, D.W., Scully, K.M., Simmons, D.M., Wegner, M., Chu, K.T.,
Swanson, L.W. and Rosenfeld, M.G.
TFE, a transcription factor expressed specifically in the anterior
pituitary during embryogenesis, defines a new class of leucine
zipper proteins
Genes Dev. 5 (10), 1739-1753 (1991)
92009166
1916262
PUBMED
REMARK
GenBank staff at the National Library of Medicine created this
entry [NCBI gibbsq 58745] from the original journal article.

FEATURES
source
1. .817
/organism="Rattus sp."
/mol_type="mRNA"
/db_xref="taxon:10118"
gene
1. .817
/gene="thyrotroph embryonic factor, TEF"
CDS
32. .817
/gene="thyrotroph embryonic factor, TEF"
/note="leucine zipper transcription factor; TEF"
/codon_start=1
/product="thyrotroph embryonic factor"
/protein_id="AAB20032.1"
/db_xref="GI:237085"
translation="MENPRETRLDKEKGKLEDEEASAAATMAVSASLMPPIWDKT
IPYDGSFHLGYMDLDFLENGIPASPTHLAQNLLLPVALEGEKESASSTASPPSS
STAIPOSETVSTESSLEKERETPSIDPNCVVDVNFNPDADIVLSSVPGGELFN
PRKHFAEDLKQPMIKKAKVFPVDEOKDEKYWTRRKNVAAKRSRDARLKENQ
ITIRAPLEKENTALRTEVAELRKEVGKCTIVSKYETKYGPL"

ORIGIN
Alignment Scores:
Pred. No.: 0.00242 Length: 817
Score: 11.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 25.58% Indels: 0
DB: 10 Gaps: 0

US-10-624-218-4 (1-43) x S58745 (1-817)

Qy 12 AsnThrAlaLeuArgThrGluValAlaGluLeu 22
|||||
Db 719 AACACAGCCTCGCGAGGAGTTGCTGAGCTT 751

RESULT 3
HSU06935
LOCUS
DEFINITION
Human thyrotroph embryonic factor (TEF) mRNA, complete cds.
ACCESSION
U06935

U06935.1 GI:606796
KEYWORDS
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Khatib, Z.A., Inaba, T., Valentine, M. and Look, A.T.
Chromosomal localization and cDNA cloning of the human DBP and TEF
genes
Genomics 23 (2), 344-351 (1994)
JOURNAL
MEDLINE
95137580
PUBMED
REFERENCE
2 (bases 1 to 855)
AUTHORS
Inaba, T.
TITLE
Direct Submission
JOURNAL
Submitted (18-FEB-1994) Toshiya Inaba, St. Jude Children's Research
Hospital, 332 North Laderdale, Memphis, TN 38105-0318, USA
FEATURES
source
1. .855
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cell_line="UOC-B1"
/cell_type="B-cell precursor"
gene
1. .855
/gene="TEF"
5'UTR
41. 50
/gene="TEF"
CDS
51. .836
/gene="TEF"
/function="transcription factor"
/codon_start=1
/evidence="experimental"
/product="thyrotroph embryonic factor"
/protein_id="AAA81373.1"
/db_xref="GI:606797"
translation="MENPREARLDDEKGKLEDEEASAAATMAVSASLMPPIWDKT
IPYDGSFHLGYMDLDFLENGIPASPTHLAQNLLLPVALEGEKESASSTASPPSS
STAIPOSETVSTESSLEKERETPSIDPNCVVDVNFNPDADIVLSSVPGGELFN
PRKHFAEDLKQPMIKKAKVFPVDEOKDEKYWTRRKNVAAKRSRDARLKENQ
ITIRAPLEKENTALRTEVAELRKEVGKCTIVSKYETKYGPL"
837. .855
/gene="TEF"
3'UTR
ORIGIN
Alignment Scores:
Pred. No.: 0.00253 Length: 855
Score: 11.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 25.58% Indels: 0
DB: 9 Gaps: 0

US-10-624-218-4 (1-43) x HSU06935 (1-855)

Qy 12 AsnThrAlaLeuArgThrGluValAlaGluLeu 22
|||||
Db 738 AACACAGCCTCGCGAGGAGTTGCCGAGCTA 770

RESULT 4
CR541827
LOCUS
DEFINITION
Homo sapiens full open reading frame cDNA clone RZPD0834D0532D for
gene TEF, thyrotrophic embryonic factor; complete cds, without
stopcodon.
ACCESSION
CR541827.1 GI:49456610
VERSION
KEYWORDS
Full ORF shuttle clone,
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 909)
 AUTHORS Halleck,A., Ebert,L., Mfoundinya,M., Schick,M., Eisenstein,S., Neubert,P., Kstrang,K., Schatten,R., Shen,B., Henze,S., Mar,W., Korn,B., Zuo,D., Hu,Y. and Labaer,J.
 TITLE Cloning of human full open reading frames in Gateway(TM) system entry vector (pDONR201)
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 909)
 AUTHORS Halleck,A., Ebert,L., Mfoundinya,M., Schick,M., Eisenstein,S., Neubert,P., Kstrang,K., Schatten,R., Shen,B., Henze,S., Mar,W., Korn,B., Zuo,D., Hu,Y. and Labaer,J.
 TITLE Direct Submission
 JOURNAL Submitted (28-JUN-2004) RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH, Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
 COMMENT RZPD: RZPD0834D0532D, ORFNo 3756
 www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=RZPD0834D0532D RZPDLIB;
 Human Full ORF Clones Gateway(TM) - RZPD (kan-resist.) RZPD LIB No. 834
 www.rzpd.de/cgi-bin/products/showLib.pl.cgi?response?libNo=834
 Contact: Inge Arlart
 RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH,
 Heubnerweg 6, D-14059 Berlin, Germany
 Tel: +49 30 32639 100
 Fax: +49 30 32639 111
 www.rzpd.de
 This clone is available from RZPD;
 Contact RZPD (customer.service@rzpd.de) for further information.
 Clone name at Harvard Institute of Proteomics
 (www.hip.harvard.edu): FLH30899.01L
 This CDS clone is part of a collection of human full ORF clones jointly established and verified by the Harvard Institute of Proteomics (HIP) and RZPD.
 This CDS has been cloned without stopcodon.
 The CDS has been inserted into pDONR201 via a BP Clonase(TM) reaction. Additional sequence has been added in front of the start codon: att. .AAAAA GCA GGC TCC ACC (ATG).
 The last codon is followed by the 3' att site: GACCCAGCTTTCTT. .att
 The clone is validated by full sequence check.
 Compared to the reference sequence NM_003216 (GI:34486096) we did not find any amino acid exchanges.
 Clone distribution: http://www.rzpd.de/products/orfclones/.
 FEATURES
 source
 1. .909
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clones="RZPD0834D0532D"
 /clone_libs="Human Full ORF Clones Gateway(TM) - RZPD"
 /lab_host="DH5alpha"
 /note="vector: pDONR201, Site_1: attP1; Site_2: attP2"
 gene
 1. .909
 /gene="TEF"
 CDS
 1. .>909
 /gene="TEF"
 /codon_start=1
 /protein_id="CAG46626.1"
 /db_xref="GI:49456611"
 /translation="MSDAGGKKPPVDPQAGPGPGRGAAGRGSLSSFPVLVKKLME
 NPPREARLDKEKKEKLEDEAAASLMPPVWDKTIPIYDGSFHLVYMDLD
 EFLLENGIPASPTLHNLHPVALEKESASSTASPPSSSTAIFQPSSTVSTES
 SLEKERTPSPIDNCVVDVNFDPADLVLSVPGGELFNPGRKHFABEDLKPPQM
 IKKAKVFVPEQDEKQKWTTRKNNVAKESRDARLKENQITIRAAFLKENTALL
 TEVAELRKEVGKCKTIVSKYKGYPL"
 Alignment Scores:
 Pred. No.: 0.00268 Length: 909
 Score: 11.00 Matches: 11
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 25.58% Indels: 0

DB: 9 Gaps: 0
 US-10-624-218-4 (1-43) x CR541827 (1-909)
 Qy 12 AsnThrAlaLeuArgThrGluValAlaGluLeu 22
 |||||
 Db 814 AACACAGCCCTGCGACGGAGGTGCGCGAGCTA 846
 RESULT 5
 CR456592 974 bp mRNA linear PRI 25-MAY-2004
 LOCUS CR456592
 DEFINITION Homo sapiens TEF full length open reading frame (ORF) cDNA clone
 (CDNA Clone C220RF:pGEM.TEF).
 ACCESSION CR456592
 VERSION CR456592.1 GI:47678714
 KEYWORDS CDNA; chromosome 22; ORF.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 974)
 AUTHORS Collins,J.E., Wright,C.L., Edwards,C.A., Davis,M.P., Grinham,J.A.,
 Cole,C.G., Goward,M.E., Aguado,B., Malliya,M., Mokrab,Y.,
 Huckle,B.J., Beare,D.M. and Dunham,I.
 TITLE Direct Submission
 JOURNAL Submitted (24-MAY-2004) Sanger Institute, Hinxton, Cambridgeshire,
 CB10 1SA, UK. E-mail enquiries: c229@sanger.ac.uk Manuscript
 Sanger Institute name: pGEM.TEF
 COMMENT Homo sapiens cDNA sequence. This sequence was generated as part of
 The Wellcome Trust Sanger Institute program to isolate cDNA clones
 representing the full length open reading frame of well annotated
 protein coding genes on human chromosome 22. For more information
 see http://www.sanger.ac.uk/HGP/Chr22/.
 FEATURES
 Location/Qualifiers
 1. .974
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /chromosome="22"
 /clones="pGEM.TEF"
 /lab_host="JM109"
 gene
 1. .974
 /gene="TEF"
 CDS
 14. .925
 /gene="TEF"
 /codon_start=1
 /protein_id="CAG30478.1"
 /db_xref="GI:47678715"
 /translation="MSDAGGKKPPVDPQAGPGPGRGAAGRGSLSSFPVLVKKLME
 NPPREARLDKEKKEKLEDEAAASLMPPVWDKTIPIYDGSFHLVYMDLD
 EFLLENGIPASPTLHNLHPVALEKESASSTASPPSSSTAIFQPSSTVSTES
 SLEKERTPSPIDNCVVDVNFDPADLVLSVPGGELFNPGRKHFABEDLKPPQM
 IKKAKVFVPEQDEKQKWTTRKNNVAKESRDARLKENQITIRAAFLKENTALL
 TEVAELRKEVGKCKTIVSKYKGYPL"
 ORIGIN
 Alignment Scores:
 Pred. No.: 0.00287 Length: 974
 Score: 11.00 Matches: 11
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 25.58% Indels: 0
 DB: 9 Gaps: 0
 US-10-624-218-4 (1-43) x CR456592 (1-974)
 Qy 12 AsnThrAlaLeuArgThrGluValAlaGluLeu 22
 |||||
 Db 827 AACACAGCCCTGCGACGGAGGTGCGCGAGCTA 859
 RESULT 6
 HSU44059
 LOCUS HSU44059 1000 bp mRNA linear PRI 23-AUG-1996

DEFINITION Human thyrotroph embryonic factor (TEF) mRNA, complete cds.
ACCESSION U44059
VERSION U44059.1 GI:1399385
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1000)
Hunger, S.P., Li, S., Fall, M.Z., Naumovski, L. and Cleary, M.L.
The proto-oncogene HLF and the related basic leucine zipper protein TEF display highly similar DNA-binding and transcriptional regulatory properties
Blood 87 (11), 4607-4617 (1996)
JOURNAL MEDLINE
PUBMED 96219638
REFERENCE 2 (bases 1 to 1000)
Hunger, S.P.
Direct Submission
Submitted (29-DEC-1995) Stephen P. Hunger, Pediatrics, University of Colorado Health Sciences Center, UCHSC, Box C229, 4200 E 9th Ave, Denver, CO 80262, USA
JOURNAL LOCATION/Qualifiers
FEATURES
source
1..1000
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="22"
/map="22q13.2"
/cell_line="ALL cell line HB11;19"
1..1000
/gene="TEF"
/genes="TEF"
/note="TEF; similar to product encoded by Gallus gallus vbp gene, GenBank Accession number U09221; PAR-bzip transcription factor"
/codon_start=1
/product="thyrotroph embryonic factor"
/protein_id="AA806497.1"
/db_xref="GI:1399386"
/translations="MSDAGGKPPVDPQAGPGFGRAGRLSGSPFLVLKLMENPREARLDKEKLEDEAAAASTMAVSASIMPPIMDKTIPYDGEFPHLEYMDLD EFLLENGIPASPHLAHNLPLVAELGEGKESASSTASPPSSTAIFQPSFTVSSTES SLKERETPPDPNCVEVDVNPADVLVSSVGGELFNP RKHKEAEDLKPPQM IKKAKVFDPEQDERKIYTRKKNVAARKRDARLKENQITIRAAFLKENTAIR TEVAELRKEVGKCKTIVSKYETKGPL"
ORIGIN
Alignment Scores:
Pred. No.: 0.00294 Length: 1000
Score: 11.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 25.58% Indels: 0
DB: 9 Gaps: 0
US-10-624-218-4 (1-43) x HSU44059 (1-1000)
Qy 12 AsnThrAlaLeuArgThrGluValAlaGluLeu 22
Db 850 AACACAGCCCTGCGGACGAGGTGGCCGAGCTA 882
RESULT 7
CQ730263
LOCUS
DEFINITION Sequence 16197 from Patent WO02068579.
ACCESSION CQ730263
VERSION CQ730263.1 GI:42303641
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
Kits, such as nucleic acid arrays, comprising a majority of humanexons or transcripts, for detecting expression and other uses thereof
Patent: WO 02068579-A 16197 06-SEP-2002;
PE Corporation (NY) (US)
JOURNAL LOCATION/Qualifiers
FEATURES
source
1..1000
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 0.00295 Length: 1003
Score: 11.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 25.58% Indels: 0
DB: 6 Gaps: 0
US-10-624-218-4 (1-43) x CQ730263 (1-1003)
Qy 12 AsnThrAlaLeuArgThrGluValAlaGluLeu 22
Db 853 AACACAGCCCTGCGGACGAGGTGGCCGAGCTA 885
RESULT 8
GGU09221 1153 bp mRNA linear VRT 30-NOV-1995
LOCUS Gallus gallus White Leghorn beta/beta isoform of vitellogenin
DEFINITION binding protein (vbp) mRNA, complete cds.
ACCESSION U09221
VERSION U09221.1 GI:483937
KEYWORDS
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
REFERENCE 1 (bases 1 to 1153)
Burch, J.B. and Davis, D.L.
Alternative promoter usage and splicing options result in the differential expression of mRNAs encoding four isoforms of chicken vitellogenin, a member of the PAR subfamily of bZIP transcription factors
Nucleic Acids Res. 22 (22), 4733-4741 (1994)
JOURNAL MEDLINE
PUBMED 95075656
REFERENCE 2 (bases 1 to 1153)
Burch, J.B.
Direct Submission
Submitted (26-APR-1994) John B.B. Burch, Fox Chase Cancer Center, 7701 Burholme Ave., Philadelphia, PA 19111, USA
JOURNAL LOCATION/Qualifiers
FEATURES
source
1..1153
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="White Leghorn"
/db_xref="taxon:9031"
/cell_type="fibroblast"
/clone_lib="Vennstrom"
/dev_stage="embryonic day 10"
1..1153
/gene="vbp"
/codon_start=1
/product="vitellogenin binding protein (VBP), beta/beta isoform"
/protein_id="AAA82156.1"
/db_xref="GI:483938"

/translation="MSVCNSAGGPAALDPFEVLKSLLEYSLPWTWKMDKEKKIKLE
DEAAATMAVSASLMPPIWDKTIYPDGESPHLEMDLDFLLNGPSSPFLDLNQ
NPLMPVAKLEKEPASATGSPVSSSTAVYQOEEAASSTESPQNERPTSPIDPDC
VEVFNPNPDADVLVSPGGLFNPGRKHFTEDLKPOMIKKAKVFPDQKDE
KYWTRKKNVAAKRSRDLKENQITIRAFLEKENTALTAEVLEIRKEVGRCKNI
VSKYTRYGPDLSDSE"

ORIGIN

Alignment Scores:

Pred. No.: 0.00338 Length: 1153
Score: 11.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 25.58% Indels: 0
DB: 5 Gaps: 0

US-10-624-218-4 (1-43) x GGU09221 (1-1153)

Oy 12 AsnThrAlaLeuArgThrGluValAlaGluLeu 22

Db 892 AATACGGCCCTGAGGACGGAGTTGCAGAGCTG 924

RESULT 9

AF194420

LOCUS

DEFINITION Mus musculus thyrotroph embryonic factor (Tef) mRNA, complete cds.

ACCESSION

AF194420

VERSION

AF194420.1

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

Location/Qualifiers

1. .1304

/organism="Mus musculus"

/mol_type="mRNA"

/strain="BALB/c"

/db_xref="taxon:10090"

/sex="male"

/tissue type="brain"

/dev stage="adult"

1. .1304

/gene="Tef"

91. .996

/note="Tef; PAR-bZIP transcription factor"

/codon start=1

/product="thyrotroph embryonic factor"

/protein_id="AAF0830.1"

/db_xref="GI:7939636"

/translation="MSDAGGGKKPPFVQAGPGPGRAGGLSGSPFLVLKLMENP

PRETRLDKEKKEKLEESAASMTVASLMPPIWDKTIYPDGESPHLEMDLDF

LLENGIPASPHLAQNLLPVALEGKESAGSSTASPPSSSTAIPOPTVTSSTSSLL

EKERTPSIDPSCVEVDWNPDPADLVLSVFGLEENPRKRFASDELKPOPMIK

KAKVFPVDEQDEKIYTRKKNVAAKRSRDLKENQITIRAFLEKENTALTAE

VLEIRKEVGRCKNIYVSKYTRYGPDLSDSE"

ORIGIN

Alignment Scores:

Pred. No.: 0.00381

Score: 11.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 25.58%

DB: 10

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

US-10-624-218-4

(1-43) x

AF194420

(1-1304)

Oy

12

AsnThrAlaLeuArgThrGluValAlaGluLeu

22

Db

898

AACACAGCCCTGCGGACGAGTTGCCGAGCTT

930

RESULT 10

AK091916

LOCUS

DEFINITION

Homo sapiens

cdna FLJ34597 fis, clone KIDNE2009367, highly similar

to THYROTROPH EMBRYONIC FACTOR.

AK091916

VERSION

AK091916.1

KEYWORDS

oligo capping; fis (full insert sequence).

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

Submitted (04-JUL-2002) Takao Isogai, FUJ Project (HRI Team); 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986) NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.

FEATURES
source
1. .2823
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="KIDNE2009367"
/tissue_type="kidney"
/clone_lib="KIDNE2"
/note="cloning vector: pME18SFL3"

ORIGIN
Alignment Scores:
Pred. No.: 0.00811 Length: 2823
Score: 11.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 25.58% Indels: 0
DB: 9 Gaps: 0

US-10-624-218-4 (1-43) x AK091916 (1-2823)

QY 12 AsnThrAlaLeuArgThrGluValAlaGluLeu 22
|||||
Db 802 AACACAGCCTCGGACGAGGTGCCGAGCTA 834
|||||

RESULT 11
AY540632
LOCUS 3103 bp mRNA linear ROD 12-APR-2004
DEFINITION Mus musculus thyrotroph embryonic factor beta isoform (Tef) mRNA, complete cds, alternatively spliced.
ACCESSION AY540632
VERSION AY540632.1 GI:42768795
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE Zhou,J., Hoggatt,A.M. and Herring,B.P.
AUTHORS (bases 1 to 3103)
TITLE Activation of the smooth muscle-specific telokin gene by thyrotroph embryonic factor (TEF)
J. Biol. Chem. 279 (16), 15929-15937 (2004)
PUBMED 14702338
REFERENCE 2 (bases 1 to 3103)
AUTHORS Zhou,J., Hoggatt,A.M. and Herring,B.P.
TITLE Direct Submission
JOURNAL Submitted (30-JAN-2004) Cellular and Integrative Physiology, Indiana University School of Medicine, 635 Barnhill Drive, MS2067, Indianapolis, IN 46202, USA
FEATURES
source
1. .3103
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/chromosome="15"
/tissue_lib="bladder"
1. .3103
/gene="Tef"
6. .863
/genes="Tef"
/notes="transcription factor; alternatively spliced"
/codon_start=1

Submitted (04-JUL-2002) Takao Isogai, FUJ Project (HRI Team); 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986) NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.

FEATURES
source
1. .2823
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="KIDNE2009367"
/tissue_type="kidney"
/clone_lib="KIDNE2"
/note="cloning vector: pME18SFL3"

ORIGIN
Alignment Scores:
Pred. No.: 0.00811 Length: 2823
Score: 11.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 25.58% Indels: 0
DB: 9 Gaps: 0

US-10-624-218-4 (1-43) x AK091916 (1-2823)

QY 12 AsnThrAlaLeuArgThrGluValAlaGluLeu 22
|||||
Db 802 AACACAGCCTCGGACGAGGTGCCGAGCTA 834
|||||

RESULT 11
AY540632
LOCUS 3103 bp mRNA linear ROD 12-APR-2004
DEFINITION Mus musculus thyrotroph embryonic factor beta isoform (Tef) mRNA, complete cds, alternatively spliced.
ACCESSION AY540632
VERSION AY540632.1 GI:42768795
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE Zhou,J., Hoggatt,A.M. and Herring,B.P.
AUTHORS (bases 1 to 3103)
TITLE Activation of the smooth muscle-specific telokin gene by thyrotroph embryonic factor (TEF)
J. Biol. Chem. 279 (16), 15929-15937 (2004)
PUBMED 14702338
REFERENCE 2 (bases 1 to 3103)
AUTHORS Zhou,J., Hoggatt,A.M. and Herring,B.P.
TITLE Direct Submission
JOURNAL Submitted (30-JAN-2004) Cellular and Integrative Physiology, Indiana University School of Medicine, 635 Barnhill Drive, MS2067, Indianapolis, IN 46202, USA
FEATURES
source
1. .3103
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/chromosome="15"
/tissue_lib="bladder"
1. .3103
/gene="Tef"
6. .863
/genes="Tef"
/notes="transcription factor; alternatively spliced"
/codon_start=1

Submitted (04-JUL-2002) Takao Isogai, FUJ Project (HRI Team); 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986) NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.

FEATURES
source
1. .2823
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="KIDNE2009367"
/tissue_type="kidney"
/clone_lib="KIDNE2"
/note="cloning vector: pME18SFL3"

ORIGIN
Alignment Scores:
Pred. No.: 0.00811 Length: 2823
Score: 11.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 25.58% Indels: 0
DB: 9 Gaps: 0

US-10-624-218-4 (1-43) x AK091916 (1-2823)

QY 12 AsnThrAlaLeuArgThrGluValAlaGluLeu 22
|||||
Db 802 AACACAGCCTCGGACGAGGTGCCGAGCTA 834
|||||

RESULT 11
AY540632
LOCUS 3103 bp mRNA linear ROD 12-APR-2004
DEFINITION Mus musculus thyrotroph embryonic factor beta isoform (Tef) mRNA, complete cds, alternatively spliced.
ACCESSION AY540632
VERSION AY540632.1 GI:42768795
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE Zhou,J., Hoggatt,A.M. and Herring,B.P.
AUTHORS (bases 1 to 3103)
TITLE Activation of the smooth muscle-specific telokin gene by thyrotroph embryonic factor (TEF)
J. Biol. Chem. 279 (16), 15929-15937 (2004)
PUBMED 14702338
REFERENCE 2 (bases 1 to 3103)
AUTHORS Zhou,J., Hoggatt,A.M. and Herring,B.P.
TITLE Direct Submission
JOURNAL Submitted (30-JAN-2004) Cellular and Integrative Physiology, Indiana University School of Medicine, 635 Barnhill Drive, MS2067, Indianapolis, IN 46202, USA
FEATURES
source
1. .3103
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/chromosome="15"
/tissue_lib="bladder"
1. .3103
/gene="Tef"
6. .863
/genes="Tef"
/notes="transcription factor; alternatively spliced"
/codon_start=1

Submitted (04-JUL-2002) Takao Isogai, FUJ Project (HRI Team); 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986) NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.

FEATURES
source
1. .2823
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="KIDNE2009367"
/tissue_type="kidney"
/clone_lib="KIDNE2"
/note="cloning vector: pME18SFL3"

ORIGIN
Alignment Scores:
Pred. No.: 0.00811 Length: 2823
Score: 11.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 25.58% Indels: 0
DB: 9 Gaps: 0

US-10-624-218-4 (1-43) x AK091916 (1-2823)

QY 12 AsnThrAlaLeuArgThrGluValAlaGluLeu 22
|||||
Db 802 AACACAGCCTCGGACGAGGTGCCGAGCTA 834
|||||

RESULT 11
AY540632
LOCUS 3103 bp mRNA linear ROD 12-APR-2004
DEFINITION Mus musculus thyrotroph embryonic factor beta isoform (Tef) mRNA, complete cds, alternatively spliced.
ACCESSION AY540632
VERSION AY540632.1 GI:42768795
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE Zhou,J., Hoggatt,A.M. and Herring,B.P.
AUTHORS (bases 1 to 3103)
TITLE Activation of the smooth muscle-specific telokin gene by thyrotroph embryonic factor (TEF)
J. Biol. Chem. 279 (16), 15929-15937 (2004)
PUBMED 14702338
REFERENCE 2 (bases 1 to 3103)
AUTHORS Zhou,J., Hoggatt,A.M. and Herring,B.P.
TITLE Direct Submission
JOURNAL Submitted (30-JAN-2004) Cellular and Integr

polyA_signal
polyA_site

ORIGIN

Alignment Scores:

Pred. No.: 0.0109 Length: 3808
Score: 11.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 25.58% Indels: 0
DB: 9 Gaps: 0

US-10-624-218-4 (1-43) x HSM805945 (1-3808)

Qy 12 AenThrAlaLeuArgThrGluValAlaGluLeu 22

Db 327 AACACAGCCCTGCGGACGAGGTGCGGAGCTA 359

RESULT 13

BC017689 4018 bp mRNA linear ROD 30-JUN-2004
LOCUS Mus musculus thymotroph embryonic factor, transcript variant 2,
DEFINITION mRNA (cDNA clone MGC:19233 IMAGE:4242534), complete cds.
ACCESSION BC017689
VERSION BC017689.1 GI:17389268
KEYWORDS MGC.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE
AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S., Loquellano, N.A., Peters, G.J.,
Abrahamson, R.D., Mullah, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettner, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smailus, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

2 (bases 1 to 4018)

Strausberg, R.

Direct Submission

Submitted (03-DEC-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgi.nci.nih.gov>
Contact: MGC help desk
Email: cgabs@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LILNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-shgc.stanford.edu>
Contact: (Dickson, Mark) mcd@paxil.stanford.edu

TITLE

JOURNAL

PUBLISHED

REFERENCE

AUTHORS

TITLE

JOURNAL

2 (bases 1 to 4018)

Strausberg, R.

Direct Submission

Submitted (03-DEC-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgi.nci.nih.gov>
Contact: MGC help desk
Email: cgabs@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LILNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-shgc.stanford.edu>
Contact: (Dickson, Mark) mcd@paxil.stanford.edu

Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, J., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LILNL at: <http://image.llnl.gov>
Series: IRAC Plate: 25 Row: j Column: 24
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 23821037.

FEATURES

source

Location/Qualifiers
1..4018
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="MGC:19233 IMAGE:4242534"
/tissue_type="Kidney, normal. 5 month old male mouse."
/clone_lib="NCI CGAP_Kid14"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
1..4018
/gene="Tef"
/db_xref="LocusID:21685"
4..825
/db_xref="MGI:98663"
/gene="Tef"
/codon_start=1
/product="thymotroph embryonic factor, isoform 2"
/protein_id="AAH17689.1"
/db_xref="GI:17389269"
/db_xref="LocusID:21685"
/db_xref="MGI:98663"
/translation="MDMPVLKSLHSLPWEKKADKEKKEKLEDESAASTMAV
SASLMPIDWKTIPTDGESEHLYMDLDFLENGIPASPTHLAQNLLLPVALEKGE
SASSTASPSSSTAIPQSTVSSTSELEKERTSPDSCVEVDVNFNPADL
VLSSVPGSELFNPKRHFPAEDLKQPMKAKKVFVDEOKDEKYTRKKNNVAAK
RSRDARKENQITIRAFLEKENTARTEVALRKVGKCKTIVSKYTKYGPL"

gene

1..4018
/gene="Tef"
/db_xref="LocusID:21685"
4..825
/db_xref="MGI:98663"
/gene="Tef"
/codon_start=1
/product="thymotroph embryonic factor, isoform 2"
/protein_id="AAH17689.1"
/db_xref="GI:17389269"
/db_xref="LocusID:21685"
/db_xref="MGI:98663"

CDS

1..4018
/gene="Tef"
/db_xref="LocusID:21685"
4..825
/db_xref="MGI:98663"
/gene="Tef"
/codon_start=1
/product="thymotroph embryonic factor, isoform 2"
/protein_id="AAH17689.1"
/db_xref="GI:17389269"
/db_xref="LocusID:21685"
/db_xref="MGI:98663"

1..4018
/gene="Tef"
/db_xref="LocusID:21685"
4..825
/db_xref="MGI:98663"
/gene="Tef"
/codon_start=1
/product="thymotroph embryonic factor, isoform 2"
/protein_id="AAH17689.1"
/db_xref="GI:17389269"
/db_xref="LocusID:21685"
/db_xref="MGI:98663"

1..4018
/gene="Tef"
/db_xref="LocusID:21685"
4..825
/db_xref="MGI:98663"
/gene="Tef"
/codon_start=1
/product="thymotroph embryonic factor, isoform 2"
/protein_id="AAH17689.1"
/db_xref="GI:17389269"
/db_xref="LocusID:21685"
/db_xref="MGI:98663"

ORIGIN

Alignment Scores:
Pred. No.: 0.0114 Length: 4018
Score: 11.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 25.58% Indels: 0
DB: 9 Gaps: 0

US-10-624-218-4 (1-43) x BC017689 (1-4018)

Qy 12 AenThrAlaLeuArgThrGluValAlaGluLeu 22

Db 727 AACACAGCCCTGCGGACGAGGTGCGGAGCTT 759

RESULT 14

AY540631

LOCUS

DEFINITION

AY540631

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBLISHED

REFERENCE

AUTHORS

TITLE

2 (bases 1 to 4064)

Zhou, J., Hoggatt, A.M. and Herring, B.P.

Activation of the smooth muscle-specific telokin gene by thymotroph
embryonic factor (TEF)

J. Biol. Chem. 279 (16), 15929-15937 (2004)

14702338

2 (bases 1 to 4064)

Zhou, J., Hoggatt, A.M. and Herring, B.P.

Direct Submission

2 (bases 1 to 4064)

Zhou, J., Hoggatt, A.M. and Herring, B.P.

Direct Submission

JOURNAL	Submitted (30-JAN-2004) Cellular and Integrative Physiology, Indiana University School of Medicine, 635 Barnhill Drive, MS2067, Indianapolis, IN 46202, USA			JOURNAL	human and mouse cDNA sequences Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)					
FEATURES	Location/Qualifiers			REFERENCE	2 (bases 1 to 4222)					
source	1. .4064 /organism="Mus musculus" /mol_type="mRNA" /db_xref="taxon:10090" /chromosome="15" /tissue_lib="bladder"			AUTHORS	Strausberg,R.					
gene	1. .4064 /gene="Tef"			TITLE	Direct Submission					
CDS	143..1048 /gene="Tef" /note="transcription factor; alternatively spliced" /codon_start=1 /product="thyrotroph embryonic factor alpha isoform" /protein_id="AAS45599.1" /db_xref="GI:42768794" /translations="MSDAGGKKPPVEPQAGPGRAAGRGSLGSFPLVLKLMENP /translation="MSDAGGKKPPVEPQAGPGRAAGRGSLGSFPLVLKLMENP PRETRLDKEKLEDESAASTWAVSASLMPPIDWKTIPYDGSFHLVYMDLDEF LLENGIPASPTHQAQNLLPVAFLEGKESASSSTASPPSSSTAIFQPSFTVSTESSL EKERTSPIDPSCEVDVNFDPADLVSSVPGGELFNPKRHRFAEEDLKQPQMIK KAKKVFPEQDEKDYWTRKKNNVAAKSRDARRLKENQITIRAAFLKENTALTETE VAELRKEVGKCKTIVSKYETKYGPL"			JOURNAL	Submitted (23-AUG-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA					
ORIGIN	Contact: nisc.mgc@nih.gov Web site: http://www.nisc.nih.gov/ Contact: nisc.mgc@nih.gov Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B., Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S., Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P., Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaapi,R., Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C., McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W., Tsurgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L., Young,A., Zhang,L.-H. and Green,E.D.			REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov					
Alignment Scores:	0.0116	Length: 4064		COMMENT	Contact: MGC help desk Email: cgapbs-r@mail.nih.gov Tissue Procurement: Jeffrey Green M.D. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland; Web site: http://www.nisc.nih.gov/					
Pred. No.:	11.00	Matches: 11		Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov						
Score:	100.00%	Conservative: 0		Series: IRAK Plate: 81 Row: 9 Column: 3						
Best Local Similarity:	100.00%	Mismatches: 0		This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 23821034.						
Query Match:	25.58%	Indels: 0		Location/Qualifiers						
DB:	10	Gaps: 0		1. .4222 /organism="Mus musculus" /mol_type="mRNA" /strain="FVB/N" /db_xref="taxon:10090" /clone="MGC:46858 IMAGE:4976241" /tissue_type="Mammary tumor. C3(1)-Tag model. Infiltrating ductal carcinoma. 5 month old virgin mouse." /clone_lib="NCI CGAP_Mam6" /lab_host="DH10B" /note="Vector: pCMV-SPORT6"						
US-10-624-218-4 (1-43) x AY540631 (1-4064)				1. .4222 /gene="Tef" /db_xref="LocusID:21685" /db_xref="MGI:98663" 121_..1026 /gene="Tef" /codon_start=1 /product="thyrotroph embryonic factor, isoform 1" /protein_id="AAH36982.1" /db_xref="GI:22477948" /db_xref="LocusID:21685" /db_xref="MGI:98663" /translation="MSDAGGKKPPVEPQAGPGRAAGRGSLGSFPLVLKLMENP PRETRLDKEKLEDESAASTWAVSASLMPPIDWKTIPYDGSFHLVYMDLDEF LLENGIPASPTHQAQNLLPVAFLEGKESASSSTASPPSSSTAIFQPSFTVSTESSL EKERTSPIDPSCEVDVNFDPADLVSSVPGGELFNPKRHRFAEEDLKQPQMIK KAKKVFPEQDEKDYWTRKKNNVAAKSRDARRLKENQITIRAAFLKENTALTETE VAELRKEVGKCKTIVSKYETKYGPL"						
QY	12	AsnThrAlaLeuArgThrGluValAlaGluLeu	22	ORIGIN						
Db	950	AAACACAGCCCTGCGAGCGAGGTTCGCAGCTT	982	Alignment Scores:						
RESULT 15				Pred. No.: 0.012 Length: 4222						
BC036982				Score: 11.00 Matches: 11						
LOCUS				Percent Similarity: 100.00% Conservative: 0						
DEFINITION	Mus musculus thyrotroph embryonic factor, transcript variant 1, mRNA (CDNA clone MGC:46858 IMAGE:4976241), complete cds.			Percent Similarity: 100.00% Mismatches: 0						
ACCESSION	BC036982			Best Local Similarity: 100.00% Indels: 0						
VERSION	BC036982.1			Query Match: 25.58% Gaps: 0						
KEYWORDS	MGC.			ORIGIN						
SOURCE	Mus musculus (house mouse)			Alignment Scores:						
ORGANISM	Mus musculus			Pred. No.: 0.012 Length: 4222						
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 4222)			Score: 11.00 Matches: 11						
AUTHORS	Krausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shennen,C.M., Schuller,G.D., Altschul,S.F., Zeeberg,B., Buettow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kettelman,M., Madan,A., Rodrigues,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalios,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.			Percent Similarity: 100.00% Mismatches: 0						
TITLE	Generation and initial analysis of more than 15,000 full-length			Best Local Similarity: 100.00% Indels: 0						

DB: 10 Gaps: 0

US-10-624-218-4 (1-43) x HC036982 (1-4222)

Qy 12 AsnThrAlaLeuArgThrGluValAlaGluLeu 22

Db 928 AACACAGCCCTGCCGACGAGGTTGCCGAGCTT 960

Search completed: July 27, 2005, 13:56:21
Job time : 1786.92 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 27, 2005, 10:05:45 ; Search time 344.597 Seconds
(without alignments)
738.685 Million cell updates/sec

Title: US-10-624-218-4
Perfect score: 43
Sequence: 1 LEIRAAFLQRNTALRTEVA.....QEVORLENSQYTRYGPL 43

Scoring table:

OLIGO	
Xgapop 60.0 , Xgapext 60.0	
Ygapop 60.0 , Ygapext 60.0	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 4330206 seqs, 2959870667 residues

Word size: 1

Total number of hits satisfying chosen parameters: 8774676

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters: -DEV=xlp
-MODE=frame+p2n.model -p/US10624218/runat 26072005 121435 3677/app query.fasta_1.796
-Q=/cgn2_1/USPTO_spool_p/US10624218/runat 26072005 121435 3677/app query.fasta_1.796
-DB=N Geneseq_16Dec04 -QFT=fastap -SUFFIX=olip2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFWT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10624218 @C@N_1_1_1241 @runat 26072005 121435 3677 -NCPU=6 -ICPU=3
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :				N Geneseq_16Dec04:*			
1: Geneseqn1980s:*				ABI99561			
2: Geneseqn1990s:*				ID ABI99561 standard; cDNA; 446 BP.			
3: Geneseqn2000s:*				AC ABI99561;			
4: Geneseqn2001as:*				DT 07-MAR-2002 (first entry)			
5: Geneseqn2001bs:*				XX Mouse ischaemic condition related cDNA sequence SEQ ID NO:564.			
6: Geneseqn2002as:*				DE Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;			
7: Geneseqn2002bs:*				XX KW vasospastic ischaemia; ischaemic condition; ischaemic disease; SB.			
8: Geneseqn2003as:*				OS Mus musculus.			
9: Geneseqn2003bs:*				XX PN WO200188188-A2.			
10: Geneseqn2003cs:*				XX PD 22-NOV-2001.			
11: Geneseqn2003ds:*				XX PF 18-MAY-2001; 2001WO-JP004192.			
12: Geneseqn2004as:*				XX PR 18-MAY-2000; 2000JP-00145977.			
13: Geneseqn2004bs:*				XX PA (UYNI-) UNIV NIHON SCHOOL JURIDICAL PERSON.			
				XX PI Ishikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y;			
				XX DR WPI; 2002-034733/04.			
				XX			

RESULT 1

6	11	25.6	4480	5	ABA16225	ABA16225 Human ner
7	11	25.6	5240	5	ABA16224	ABA16224 Human ner
8	9	20.9	192	3	AAC06163	AAC06163 Human sec
9	9	20.9	475	9	ACH25568	ACH25568 Human adu
10	9	20.9	499	5	ABV15119	ABV15119 Human pro
11	9	20.9	509	5	ABV23789	ABV23789 Human pro
12	9	20.9	509	5	ABV29664	ABV29664 Human pro
13	9	20.9	526	5	ABV05950	ABV05950 Human pro
14	9	20.9	617	5	ABV44983	ABV44983 Human pro
15	9	20.9	617	5	ABV35920	ABV35920 Human pro
16	9	20.9	1071	10	ADJ80238	ADJ80238 Novel hum
17	9	20.9	1507	12	ADN61958	ADN61958 Human cDN
18	9	20.9	1508	6	ABX97121	ABX97121 Human NOV
19	9	20.9	1597	6	ABK48722	ABK48722 CDNA enco
20	9	20.9	1649	4	AAK52590	AAK52590 Human pol
21	9	20.9	1657	4	AAK51606	AAK51606 Human pol
22	9	20.9	1776	6	ABX97120	ABX97120 Human NOV
23	9	20.9	1776	12	ADN61956	ADN61956 Human cDN
24	9	20.9	1776	13	ADR66191	ADR66191 Human pro
25	9	20.9	1818	10	ADAS2954	ADAS2954 Human cod
c 26	8	18.6	542	13	ACN60962	ACN60962 Cotton gy
c 27	8	18.6	557	12	ACH74272	ACH74272 Human gen
c 28	8	18.6	2808	10	ACF67604	ACF67604 Photorhab
29	8	18.6	110000	10	ACF65383_1	Continuation (2 of
30	8	18.6	110000	10	ACF67367_01	Continuation (2 of
c 31	7	16.3	168	7	ADS65379	ADS65379 Corn seed
32	7	16.3	199	6	ABL79631	ABL79631 Human ova
c 33	7	16.3	207	6	AAD30815	AAD30815 Clone G30
c 34	7	16.3	220	4	AAD15055	AAD15055 Oryctolag
c 35	7	16.3	257	6	ABL79659	ABL79659 Human ova
36	7	16.3	279	10	ACF72150	ACF72150 Photorhab
37	7	16.3	296	2	AAT84355	AAT84355 4heptad-F
38	7	16.3	313	2	AAQ60875	AAQ60875 Human bra
c 39	7	16.3	329	2	AAV88328	AAV88328 EST clone
40	7	16.3	330	6	ABN94395	ABN94395 Gene #893
41	7	16.3	388	8	ACC60048	ACC60048 Rice leaf
c 42	7	16.3	434	9	ACH44541	ACH44541 Human foe
43	7	16.3	440	6	ABL87579	ABL87579 Human ova
44	7	16.3	443	4	AAS38635	AAS38635 Novel hum
45	7	16.3	452	4	AAK89408	AAK89408 Human dig

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES						
Result No.	Score	Query Match	Length	ID	Description	
1	11	25.6	446	6	ABI99561	Mouse isc
2	11	25.6	999	13	ADQ87162	Human tum
3	11	25.6	999	13	ADQ87502	Human tum
4	11	25.6	1000	12	ADI82448	Human mod
5	11	25.6	1000	13	ACN40954	Tumour-as

PT Examining the ischemic condition (e.g. occlusive ischemia) by measuring
PT expression levels of particular genes defined in the specification or by
PT determining the expression profile of a gene group comprising these
PT genes.

XX Claim 2; Page 1498; 2690pp; English.

XX The present invention describes a method for examining ischaemic
XX conditions, comprising measuring the expression levels of particular
XX genes (1) in a test sample or determining the expression profile of a
XX gene group in the sample comprising genes selected from (1). The method
XX is useful for examining the ischaemic condition (e.g. compressive
XX ischaemia, occlusive ischaemia or vasospastic ischaemia) by measuring
XX expression levels of particular genes (ABI99202 to ABI99912, encoding the
XX protein sequences in ABBS7020 to ABBS7374) or by determining the
XX expression profile of a gene group comprising these genes. The expression
XX levels or expression profiles produced by these genes are used as an
XX indicator when screening for ischaemic condition-improving drugs or
XX therapeutics for ischaemic diseases. ABI99913 and ABI99914 represent PCR
XX primers for a mouse ischaemic condition related sequence, which are used
XX in the exemplification of the present invention

SQ Sequence 446 BP; 109 A; 127 C; 129 G; 81 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 0.0157 Length: 446
Score: 11.00 Matches: 11
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 25.58% Indels: 0
DB: 6 Gaps: 0

US-10-624-218-4 (1-43) x ABI99561 (1-446)

Qy 12 AsnThrAlaIeuArgThrGluValAlaGluIeu 22
Db 112 AACACAGCCCTGCGACGAGGTTGCGGAGCTT 144

RESULT 2

ADQ87162
ID ADQ87162 standard; cDNA; 999 BP.

XX ADQ87162;

XX 07-OCT-2004 (first entry)

XX Human tumour-associated antigenic target (TAT) cDNA sequence #4038.

XX human; tumour-associated antigenic target; TAT; cytostatic; gene therapy;
XX cancer; cell proliferative disorder; gene; ss.

XX Homo sapiens.

XX WO2004060270-A2.

XX 22-JUL-2004.

XX 15-OCT-2003; 2003WO-US029126.

XX 18-OCT-2002; 2002US-0418988P.

XX (GETH) GENENTECH INC.

XX (WUTD/) WU T. D.

XX (ZHOU/) ZHOU Y.

XX Wu TD, Zhou Y;

XX WPI; 2004-534300/51.

XX New nucleic acid molecule and encoded polypeptide, for diagnosing,
XX preventing or treating cell proliferative disorders such as cancer.

XX Claim 1; SEQ ID NO 4038; 5504pp; English.

XX The present invention describes an isolated tumour-associated antigenic
XX target (TAT) nucleic acid comprising: (a) any of 4622 nucleotide
XX sequences (see SEQ ID NO:1 to 4622); (b) the full-length coding region of
XX (a); (c) the complement of (a) or (b); (d) a sequence that has 80%
XX sequence identity to (a)-(c); or (e) a sequence that hybridises to (a)-
XX nucleic acid; (2) a host cell comprising the above expression vector; (3)
XX a process for producing a polypeptide; (4) an isolated polypeptide
XX comprising: (a) an amino acid sequence encoded by any of the above
XX nucleotide sequences; (b) an amino acid sequence encoded by the full-
XX length coding region of the above nucleotide sequences; or (c) a sequence
XX having at least 80% identical to (a) or (b); (5) a chimeric polypeptide
XX comprising the above polypeptide fused to a heterologous polypeptide; (6)
XX an isolated antibody that binds to the above polypeptide; (7) a process
XX for producing the antibody; (8) an isolated oligopeptide that binds to
XX the above polypeptide; (9) a tumour-associated antigenic target (TAT)
XX binding organic molecule that binds to the above polypeptide; (10) a
XX composition of matter comprising the above (chimeric) polypeptide,
XX antibody, oligopeptide or TAT binding organic molecule, in combination
XX with a carrier; (11) an article of manufacture comprising a container and
XX the composition of matter contained within the container; (12) methods of
XX inhibiting the growth of a cell that expresses the above protein, where
XX the growth of the cell is at least in part dependent upon a growth
XX potentiating effect of the above protein; (13) a method of
XX therapeutically treating a mammal having a cancerous tumour comprising
XX cells that express the above protein; (14) a method of determining the
XX presence of a protein in a sample suspected of containing the protein in a
XX described above; (15) methods of diagnosing the presence of a tumour in a
XX mammal; (16) a method for treating or preventing a cell proliferative
XX disorder associated with increased expression or activity of the above
XX protein; and (17) a method of binding an antibody, oligopeptide or
XX organic molecule to a cell that expresses the protein described above.
XX The TAT sequences have cytostatic activities, and can be used in gene
XX therapy. The composition and methods are useful for diagnosing,
XX preventing or treating cancer. The composition is also used for preparing
XX a medicament for the therapeutic treatment or diagnostic detection of a
XX cell proliferative disorder or cancer. The present sequence represents a
XX human TAT cDNA sequence from the present invention.

SQ Sequence 999 BP; 231 A; 311 C; 307 G; 150 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 0.0331 Length: 999
Score: 11.00 Matches: 11
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 25.58% Indels: 0
DB: 13 Gaps: 0

US-10-624-218-4 (1-43) x ADQ87162 (1-999)

Qy 12 AsnThrAlaIeuArgThrGluValAlaGluIeu 22
Db 849 AACACAGCCCTGCGACGAGGTTGCGGAGCTA 881

RESULT 3

ADQ87502
ID ADQ87502 standard; cDNA; 999 BP.

XX ADQ87502;

XX 07-OCT-2004 (first entry)

XX Human tumour-associated antigenic target (TAT) cDNA sequence #4380.

XX human; tumour-associated antigenic target; TAT; cytostatic; gene therapy;
XX cancer; cell proliferative disorder; gene; ss.

XX Homo sapiens.

XX WO2004060270-A2.

XX

PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234221P.
PR 21-SEP-2000; 2000US-0234223P.
PR 25-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 26-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235835P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 20-OCT-2000; 2000US-0242221P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.

PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250391P.
PR 01-DEC-2000; 2000US-0251160P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
PI
XX
XX WPI; 2001-541565/60.
DR
XX
XX Nucleic acids encoding 3224 human nervous system antigen polypeptides,
PT useful for preventing, diagnosing and/or treating nervous system cancers
PT and metastases.
XX
PS Disclosure; SEQ ID NO 8556; 1701pp + Sequence Listing; English.
XX
XX The invention relates to novel genes (ABA11004-ABA21534) and proteins
CC (AB14678-AB18001) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in
CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
CC cardiovascular disorders such as myocardial ischaemias; (d) wound healing
CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
CC infectious diseases such as viral, bacterial, fungal and parasitic
CC infections. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 4480 BP; 913 A; 1195 C; 1214 G; 1158 T; 0 U; 0 Other;

Alignment Scores:
Pred.No.: 0.133 Length: 4480
Score: 11.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 25.58% Indels: 0
DB: 5 Gaps: 0

US-10-624-218-4 (1-43) x ABA16225 (1-4480)

QY 12 AsnThrAlaLeuArgThrGluValAlaGluLeu 22
|||
Db 1769 AACACAGCCCTCGGACGGAGGTGGCCGAGCTA 1801

RESULT 7

ABA16224
ID ABA16224 standard; DNA; 5240 BP.
XX AC ABA16224;
XX AC ABA16224;
XX AC ABA16224;
DT DT 23-JAN-2002 (first entry)
XX DE Human nervous system related polynucleotide SEQ ID NO 8555.
XX XX
XX Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;
KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;
KW antiparkinsonian; antischlicking; antianaemic; antiarthritic; cancer;
KW antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
KW antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;
KW antiparasitic; cardiac; immune disorder; cardiovascular disorder;
KW neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.
XX XX
XX Homo sapiens.
XX XX
XX WC200159063-A2.
XX PD 16-AUG-2001.
XX PF 17-JAN-2001; 2001WO-US0001334.
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 11-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226661P.
PR 22-AUG-2000; 2000US-0226688P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 06-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 08-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0179065P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 21-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 20-OCT-2000; 2000US-0242221P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.

PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250391P.
PR 01-DEC-2000; 2000US-0251160P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-541565/60.
XX Nucleic acids encoding 3224 human nervous system antigen polypeptides,
PT useful for preventing, diagnosing and/or treating nervous system cancers
PT and metastases.
XX
XX Disclosure; SEQ ID NO 8555; 1701pp + Sequence Listing; English.
XX
XX The invention relates to novel genes (ABAI1004-ABA21534) and proteins
CC (ABB14678-ABB18001) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in
CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
CC cardiovascular disorders such as myocardial ischaemias; (d) wound healing
CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
CC infectious diseases such as viral, bacterial, fungal and parasitic
CC infections. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX SQ Sequence 5240 BP; 1079 A; 1354 C; 1452 G; 1355 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 0.154 Length: 5240
Score: 11.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 25.58% Indels: 0
DB: 5 Gaps: 0

US-10-624-218-4 (1-43) x ABA16224 (1-5240)
QY 12 AenThrAlaLeuArgThrGluValAlaGluLeu 22
DB 1769 AACACAGCCCTCGGACGAGGTCGCGAGCTA 1801
RESULT 8
AAC06163
ID AAC06163 standard; cDNA; 192 BP.
XX
XX AAC06163;
XX
XX 06-OCT-2000 (first entry)
DT
DE Human secreted protein 5' EST, SEQ ID NO: 10238.
XX
XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;

KW gene therapy; chromosome mapping; ss.
XX
XX Homo sapiens.
XX
XX EP1033401-A2.
XX
XX 06-SEP-2000.
XX
XX 21-FEB-2000; 2000EP-00200610.
XX
XX 26-FEB-1999; 99US-0122487P.
XX
XX (GEST) GENSET.
XX
XX Dumas Milne Edwards J, Duclert A, Giordano J;
XX WPI; 2000-500381/45.
XX
XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures.
XX
XX Claim 1; SEQ ID NO 10238; 71pp + Sequence Listing; English.
XX
XX The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC identified within the present sequence. The 5' ESTs were prepared from
CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
CC sequences usually correspond mainly to the 3' untranslated region (UTR)
CC of the mRNA because they are often obtained from oligo-dT primed cDNA
CC libraries. Such ESTs are not well suited for isolating cDNA sequences
CC derived from the 5' ends of mRNAs and even in those cases where longer
CC cDNA sequences have been obtained, the full 5' UTR is rarely included. 5'
CC ESTs are derived from mRNAs with intact 5' ends and can therefore be used
CC to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in
CC diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC They are used to obtain upstream regulatory sequences and to design
CC expression and secretion vectors
XX
XX SQ Sequence 192 BP; 64 A; 38 C; 51 G; 36 T; 0 U; 3 Other;

Alignment Scores:
Pred. No.: 0.946 Length: 192
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 20.93% Indels: 0
DB: 3 Gaps: 0

US-10-624-218-4 (1-43) x AAC06163 (1-192)
QY 21 GluLeuGluGlnGluValGlnArgLeu 29
DB 142 GAACCTGGAGCAGAGCTGCGAAGACTT 168
RESULT 9
ACH25568
ID ACH25568 standard; cDNA; 475 BP.
XX
XX ACH25568;
XX
XX 13-OCT-2003 (first entry)
DT
XX Human adult ovary cDNA #3948.
XX
XX Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
KW genome mapping; biodiversity; genetic disorder.
XX
XX Homo sapiens.
XX
XX US2003073623-A1.
XX
XX 17-APR-2003.

XX 30-JUL-2001; 2001US-00918995.
PF
XX
PR 30-JUL-2001; 2001US-00918995.
XX
PA (DRMA/) DRMANAC R T.
XX (LABA/) LABAT I.
PA (STAC/) STACHE-CRAIN B.
PA (DICK/) DICKSON M C.
PA (JONE/) JONES L W.
XX
XX Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;
PI WPI; 2003-615964/58.
XX
XX New polynucleotide sequences obtained from various cDNA libraries, useful
PT as hybridization probes, as oligomers for PCR, for chromosome and gene
PT mapping, in the recombinant production of protein, or in generating
PT antisense DNA or RNA.
XX
XX Claim 1; SEQ ID NO 12780; 44pp; English.
XX
XX The invention relates to an isolated polynucleotide comprising any one of
CC 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
CC determined by the technique of SSH (sequencing by hybridisation). Also
CC included is a purified polypeptide comprising a sequence corresponding to
CC a reading frame of the novel polynucleotide. The nucleic acid sequences
CC are useful in diagnostics as expressed sequence tags (EST) for
CC identifying expressed genes or for physical mapping of the human genome,
CC in forensics, in assessing biodiversity, or in identifying mutations
CC responsible for genetic disorders and other traits. The nucleotide
CC sequences are also useful as hybridisation probes, as oligomers for PCR,
CC for chromosome and gene mapping, in the recombinant production of
CC protein, or in generating antibodies specific for it. The present sequence
CC is one of the 38043 isolated cDNA/EST sequences. Note: The present sequence
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20030073623
XX
SQ Sequence 475 BP; 162 A; 95 C; 124 G; 93 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.: 2.19 Length: 475
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 20.93% Indels: 0
DB: 9 Gaps: 0

US-10-624-218-4 (1-43) x ACH25568 (1-475)
Qy 21 GluLeuGluGlnGluValGlnArgLeu 29
Db 160 GAACCTGGAGCAAGAGGTGCAAGACTT 186
RESULT 10
ABV15119
ID ABV15119 standard; cDNA; 499 BP.
XX AC
XX ABV15119;
XX
XX 13-SEP-2002 (first entry)
XX
XX Human prostate expression marker cDNA 15110.
XX
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX
XX Homo sapiens.
XX
XX WO200160860-A2.
FN
XX

PD 23-AUG-2001.
XX
XX 20-FEB-2001; 2001WO-US005171.
XX
XX 17-FEB-2000; 2000US-0183319P.
PR 16-MAR-2000; 2000US-0189862P.
XX 25-MAY-2000; 2000US-0207454P.
PR 09-JUN-2000; 2000US-0211314P.
XX 18-JUL-2000; 2000US-0219007P.
PR 13-DEC-2000; 2000US-0255281P.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
PA
XX Schlegel R, Endege WO, Monahan JE;
PI WPI; 2001-662795/76.
XX
XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX
XX Claim 1; Page 2536; 11750pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
SQ Sequence 499 BP; 171 A; 95 C; 121 G; 112 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2.29 Length: 499
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 20.93% Indels: 0
DB: 5 Gaps: 0

US-10-624-218-4 (1-43) x ABV15119 (1-499)
Qy 21 GluLeuGluGlnGluValGlnArgLeu 29
Db 5 GAACCTGGAGCAAGAGGTGCAAGACTT 31
RESULT 11
ABV23789
ID ABV23789 standard; cDNA; 509 BP.
XX AC
XX ABV23789;
XX
XX 16-SEP-2002 (first entry)
XX
XX Human prostate expression marker cDNA 23780.
XX
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX
XX Homo sapiens.
XX
XX WO200160860-A2.
FN
XX 23-AUG-2001.
XX
XX 20-FEB-2001; 2001WO-US005171.
PF
XX

```
PR 17-FEB-2000; 2000US-0183319P.
PR 16-MAR-2000; 2000US-0189862P.
PR 25-MAY-2000; 2000US-0207454P.
PR 09-JUN-2000; 2000US-0211314P.
PR 18-JUL-2000; 2000US-0219007P.
PR 13-DEC-2000; 2000US-0255281P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Schlegel R, Endege WO, Monahan JE;
XX WPI; 2001-662795/76.
XX
XX Novel isolated nucleic acid molecule associated with cancerous state of
XX prostate cells and correlating with presence of prostate cancer, useful
XX for detecting presence of prostate cancer, stage of prostate cancer.
XX
XX Claim 1; Page 4376; 11750pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule (I) comprising
XX a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
XX specification or its complement. (I) is useful for: (a) assessing whether
XX a patient is afflicted with prostate cancer; (b) monitoring the
XX progression of prostate cancer in a patient; (c) assessing the efficacy
XX of a test compound to inhibit prostate cancer in a patient; (d) assessing
XX the efficacy of a therapy for inhibiting prostate cancer in a patient;
XX (e) selecting a composition for inhibiting prostate cancer in a patient;
XX (f) assessing the prostate cell carcinogenic potential of a compound; (g)
XX determining whether prostate cancer has metastasized in a patient; (h)
XX assessing the aggressiveness or indolence of prostate cancer in a patient
XX ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
XX Sequence 509 BP; 173 A; 96 C; 123 G; 115 T; 0 U; 2 Other;
SQ
Alignment Scores:
Pred. No.: 2.33 Length: 509
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 20.93% Indels: 0
DB: 5 Gaps: 0
US-10-624-218-4 (1-43) x ABV23789 (1-509)
Qy 21 GluLeuGluGlnGluValGlnArgLeu 29
Db 5 GAACGTGGAGCAAGAGGTGCAAGACTT 31
RESULT 12
ABV29664
ID ABV29664 standard; cDNA; 509 BP.
XX
XX AC ABV29664;
XX
XX 16-SEP-2002 (first entry)
XX
XX Human prostate expression marker cDNA 29655.
XX
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX pharmacogenomic marker; gene; ss.
XX
XX Homo sapiens.
XX
XX WO200160860-A2.
XX
XX 23-AUG-2001.
XX
XX 20-FEB-2001; 2001WO-US005171.
XX
XX 17-FEB-2000; 2000US-0183319P.
XX
XX 16-MAR-2000; 2000US-0189862P.
XX
XX 25-MAY-2000; 2000US-0207454P.
XX
XX 09-JUN-2000; 2000US-0211314P.
XX
XX 18-JUL-2000; 2000US-0219007P.
XX
XX 13-DEC-2000; 2000US-0255281P.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
PA
```

```
PR 18-JUL-2000; 2000US-0219007P.
PR 13-DEC-2000; 2000US-0255281P.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Schlegel R, Endege WO, Monahan JE;
XX WPI; 2001-662795/76.
XX
XX Novel isolated nucleic acid molecule associated with cancerous state of
XX prostate cells and correlating with presence of prostate cancer, useful
XX for detecting presence of prostate cancer, stage of prostate cancer.
XX
XX Claim 1; Page 6369; 11750pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule (I) comprising
XX a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
XX specification or its complement. (I) is useful for: (a) assessing whether
XX a patient is afflicted with prostate cancer; (b) monitoring the
XX progression of prostate cancer in a patient; (c) assessing the efficacy
XX of a test compound to inhibit prostate cancer in a patient; (d) assessing
XX the efficacy of a therapy for inhibiting prostate cancer in a patient;
XX (e) selecting a composition for inhibiting prostate cancer in a patient;
XX (f) assessing the prostate cell carcinogenic potential of a compound; (g)
XX determining whether prostate cancer has metastasized in a patient; (h)
XX assessing the aggressiveness or indolence of prostate cancer in a patient
XX ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
XX Sequence 509 BP; 173 A; 96 C; 123 G; 115 T; 0 U; 2 Other;
SQ
Alignment Scores:
Pred. No.: 2.33 Length: 509
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 20.93% Indels: 0
DB: 5 Gaps: 0
US-10-624-218-4 (1-43) x ABV29664 (1-509)
Qy 21 GluLeuGluGlnGluValGlnArgLeu 29
Db 5 GAACGTGGAGCAAGAGGTGCAAGACTT 31
RESULT 13
ABV05950
ID ABV05950 standard; cDNA; 526 BP.
XX
XX AC ABV05950;
XX
XX 13-SEP-2002 (first entry)
XX
XX Human prostate expression marker cDNA 5941.
XX
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX pharmacogenomic marker; gene; ss.
XX
XX Homo sapiens.
XX
XX WO200160860-A2.
XX
XX 23-AUG-2001.
XX
XX 20-FEB-2001; 2001WO-US005171.
XX
XX 17-FEB-2000; 2000US-0183319P.
XX
XX 16-MAR-2000; 2000US-0189862P.
XX
XX 25-MAY-2000; 2000US-0207454P.
XX
XX 09-JUN-2000; 2000US-0211314P.
XX
XX 18-JUL-2000; 2000US-0219007P.
XX
XX 13-DEC-2000; 2000US-0255281P.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
PA
```

XX Schlegel R, Endege WO, Monahan JE;
PI WPI; 2001-662795/76.
XX
XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX
XX Claim 1; Page 988; 11750pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the efficacy
CC of progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
SQ Sequence 526 BP; 172 A; 105 C; 132 G; 117 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2.4 Length: 526
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 20.93% Indels: 0
DB: 5 Gaps: 0

US-10-624-218-4 (1-43) x ABV05950 (1-526)

QY 21 GluLeuGluGlnGluValGlnArgLeu 29
DB 32 GAACGGAGCAAGAGGTGCAGACTT 58

RESULT 14
ABV44983
ID ABV44983 standard; cDNA; 617 BP.
XX
AC ABV44983;
XX
XX 16-SEP-2002 (first entry)
XX Human prostate expression marker cDNA 44974.
XX
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX
XX Homo sapiens.
XX
XX WO200160860-A2.
XX
XX 23-AUG-2001.
XX
XX 20-FEB-2001; 2001WO-US005171.
XX
XX 17-FEB-2000; 2000US-0183319P.
XX
XX 16-MAR-2000; 2000US-0189862P.
XX
XX 25-MAY-2000; 2000US-0207454P.
XX
XX 09-JUN-2000; 2000US-0211314P.
XX
XX 18-JUL-2000; 2000US-0219007P.
XX
XX 13-DEC-2000; 2000US-0255281P.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Schlegel R, Endege WO, Monahan JE;
PI WPI; 2001-662795/76.
XX
XX

XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX
XX Claim 1; Page 8914; 11750pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the efficacy
CC of progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
SQ Sequence 617 BP; 194 A; 126 C; 158 G; 139 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2.79 Length: 617
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 20.93% Indels: 0
DB: 5 Gaps: 0

US-10-624-218-4 (1-43) x ABV44983 (1-617)

QY 21 GluLeuGluGlnGluValGlnArgLeu 29
DB 113 GAACGGAGCAAGAGGTGCAGACTT 139

RESULT 15
ABV35920
ID ABV35920 standard; cDNA; 617 BP.
XX
XX ABV35920;
XX
XX 16-SEP-2002 (first entry)
XX Human prostate expression marker cDNA 35911.
XX
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX
XX Homo sapiens.
XX
XX WO200160860-A2.
XX
XX 23-AUG-2001.
XX
XX 20-FEB-2001; 2001WO-US005171.
XX
XX 17-FEB-2000; 2000US-0183319P.
XX
XX 16-MAR-2000; 2000US-0189862P.
XX
XX 25-MAY-2000; 2000US-0207454P.
XX
XX 09-JUN-2000; 2000US-0211314P.
XX
XX 18-JUL-2000; 2000US-0219007P.
XX
XX 13-DEC-2000; 2000US-0255281P.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Schlegel R, Endege WO, Monahan JE;
PI WPI; 2001-662795/76.
XX
XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX

XX Claim 1; Page 7453; 11750pp; English.
PS
XX
CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker

SQ Sequence 617 BP; 194 A; 126 C; 158 G; 139 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2.79 Length: 617
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 20.93% Indels: 0
DB: 5 Gaps: 0

US-10-624-218-4 (1-43) x ABV35920 (1-617)

Oy 21 GluLeuGluGlnGluValGlnArgLeu 29
Db 113 GAACGGAGCAAGAGGTGCAAGACTT 139

Search completed: July 27, 2005, 12:14:45
Job time : 350.597 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 27, 2005, 11:44:30 ; Search time 100.333 Seconds
(without alignments)
701.261 Million cell updates/sec

Title: US-10-624-218-4
Perfect score: 43
Sequence: 1 LETRAPIRORNTALRTEVA.....QEVORLENEVSQIETRYGPL 43

Scoring table:
OLIGO Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Word size: 1

Total number of hits satisfying chosen parameters: 2398817

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=xlp
-Q=/cgn2_1/USPRO_pool_p/US10624218/runat_26072005_121436_3711/app_query.fasta_1.796
-DB=Issued Patents NA -QMT=fastap -SUFFIX=olip2n.rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10624218_QCGN 1 1 213 @runat_26072005_121436_3711 -NCPUS=3
-NO_MMAP -LARGOQUERY -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PTCUS.COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	11	25.6	1165	4	Sequence 5024, Ap
2	11	25.6	32584	4	Sequence 16766, A
3	9	20.9	192	4	Sequence 10238, A
4	9	20.9	418	4	Sequence 3746, Ap
5	8	18.6	25	4	Sequence 96364, A
6	8	18.6	145241	4	Sequence 17394, A
7	8	18.6	145241	4	Sequence 17395, A
8	7	16.3	296	2	Sequence 32, Appl
9	7	16.3	296	3	Sequence 32, Appl
10	7	16.3	601	4	Sequence 50871, A
11	7	16.3	601	4	Sequence 50872, A
12	7	16.3	601	4	Sequence 93464, A

13	7	16.3	720	3	US-09-479-309-8	Sequence 8, Appli
14	7	16.3	720	4	US-09-627-393-8	Sequence 8, Appli
15	7	16.3	770	1	US-08-428-091-3	Sequence 3, Appli
16	7	16.3	946	4	US-09-640-211A-448	Sequence 448, App
17	7	16.3	1065	4	US-08-976-063E-3	Sequence 3, Appli
18	7	16.3	1278	4	US-09-489-039A-3845	Sequence 3845, Ap
19	7	16.3	1323	4	US-09-134-000C-948	Sequence 948, App
20	7	16.3	2940	4	US-09-270-767-14350	Sequence 14350, A
21	7	16.3	3779	4	US-09-902-540-9296	Sequence 9296, Ap
22	7	16.3	11387	4	US-09-949-016-13811	Sequence 13811, A
23	7	16.3	12624	4	US-09-902-540-993	Sequence 993, App
24	7	16.3	24056	4	US-09-949-016-14320	Sequence 14320, A
25	7	16.3	32679	4	US-08-976-063E-1	Sequence 1, Appli
26	7	16.3	37254	4	US-09-949-016-15973	Sequence 15973, A
27	7	16.3	47363	4	US-09-949-016-13420	Sequence 13420, A
28	7	16.3	94987	4	US-09-949-016-12510	Sequence 12510, A
29	7	16.3	96987	4	US-09-949-016-14429	Sequence 14429, A
30	7	16.3	113701	4	US-09-949-016-13214	Sequence 13214, A
31	7	16.3	181251	4	US-09-949-016-15970	Sequence 15970, A
32	7	16.3	192506	4	US-09-949-016-15830	Sequence 15830, A
33	7	16.3	228851	4	US-09-949-016-13781	Sequence 13781, A
34	6	14.0	25	4	US-09-396-196G-3653	Sequence 3653, A
35	6	14.0	25	4	US-09-396-196G-83406	Sequence 83406, A
36	6	14.0	25	4	US-09-396-196G-96362	Sequence 96362, A
37	6	14.0	25	4	US-09-396-196G-112505	Sequence 112505, A
38	6	14.0	25	4	US-09-396-196G-112506	Sequence 112506, A
39	6	14.0	42	1	US-08-375-116A-133	Sequence 133, App
40	6	14.0	75	1	US-08-442-572-55	Sequence 55, Appl
41	6	14.0	75	1	US-08-361-795-55	Sequence 55, Appl
42	6	14.0	75	5	PCT-US95-05600-138	Sequence 138, App
43	6	14.0	117	4	US-09-902-540-7030	Sequence 7030, Ap
44	6	14.0	124	3	US-07-757-022B-5	Sequence 5, Appli
45	6	14.0	186	4	US-09-270-767-25612	Sequence 25612, A

ALIGNMENTS

RESULT 1

US-09-949-016-5024
; Sequence 5024, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949, 016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5024
; LENGTH: 1165
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-5024

Alignment Scores:
Pred. No.: 0.00783 Length: 1165
Score: 11.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 25.58% Indels: 0
DB: 4 Gaps: 0

US-10-624-218-4 (1-43) x US-09-949-016-5024 (1-1165)

Qy 12 AsnThrAlaLeuArgThrGluValAlaGluLeu 22

; LENGTH: 25
; TYPE: DNA
; ORGANISM: mus musculus
US-09-396-196G-96364
Alignment Scores:
Pred. No.: 0.374 Length: 25
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 18.60% Indels: 0
DB: 4 Gaps: 0
US-10-624-218-4 (1-43) x US-09-396-196G-96364 (1-25)
Qy 15 LeuArgThrGluValAlaGluLeu 22
Db 1 CTGCGACGGAGGTTCGCGAGCTT 24
RESULT 6
US-09-949-016-17394/c
; Sequence 17394, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17394
; LENGTH: 145241
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17394
Alignment Scores:
Pred. No.: 1.24e+03 Length: 145241
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 18.60% Indels: 0
DB: 4 Gaps: 0
US-10-624-218-4 (1-43) x US-09-949-016-17394 (1-145241)
Qy 6 AlaPheLeuArgGlnArgAsnThr 13
Db 70791 GCCTTTCTTAGCAACGAATACA 70768
RESULT 7
US-09-949-016-17395/c
; Sequence 17395, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17395
; LENGTH: 145241
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17395
Alignment Scores:
Pred. No.: 1.24e+03 Length: 145241
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 18.60% Indels: 0
DB: 4 Gaps: 0
US-10-624-218-4 (1-43) x US-09-949-016-17395 (1-145241)
Qy 6 AlaPheLeuArgGlnArgAsnThr 13
Db 70791 GCCTTTCTTAGCAACGAATACA 70768
RESULT 8
US-08-690-011A-32
; Sequence 32, Application US/08690011A
; Patent No. 5942433
; GENERAL INFORMATION:
; APPLICANT: VINSON, Charles R.
; APPLICANT: KRYLOV, Dmitry
; TITLE OF INVENTION: EXTENSION OF A PROTEIN-PROTEIN
; TITLE OF INVENTION: INTERACTION SURFACE TO INACTIVATE THE FUNCTION OF A
; TITLE OF INVENTION: CELLULAR PROTEIN
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30B
; CURRENT APPLICATION DATA: US/08/690,011A
; APPLICATION NUMBER: US/08/690,011A
; FILING DATE: 31-JUL-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/001,654
; FILING DATE: 31-JUL-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/018,496
; FILING DATE: 29-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Serunian, Leslie A.
; REGISTRATION NUMBER: 35,353
; REFERENCE/DOCKET NUMBER: 2026-4199US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)758-4800
; TELEFAX: (212)751-6849
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 296 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-690-011A-32
Alignment Scores:

Pred. No.: 45.4 Length: 296
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 16.28% Indels: 0
DB: 2 Gaps: 0

US-10-624-218-4 (1-43) x US-08-690-011A-32 (1-296)

Qy 20 AlacluLeuGluGlnGluVal 26
Db 130 GCTGAACCTCGAGCAGGAAGTG 150

RESULT 9

US-09-299-495F-32
Sequence 32, Application US/09299495F

Patent No. 6361968

GENERAL INFORMATION:

APPLICANT: VINSON, Charles R.

KRYLOV, Dmitry

TITLE OF INVENTION: INTERSECTION OF A PROTEIN-PROTEIN

INTERACTION SURFACE TO INACTIVATE THE FUNCTION OF A

CELLULAR PROTEIN

NUMBER OF SEQUENCES: 64

CORRESPONDENCE ADDRESS:

ADDRESSEE: MORGAN & FINNEGAN, L.L.P.

STREET: 345 Park Avenue

CITY: New York

STATE: NY

COUNTRY: USA

ZIP: 10154-0053

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: MS WORD 97

SOFTWARE: ASCII

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/299,495F

FILING DATE: 26-Apr-1999

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/001,654

FILING DATE: 31-JUL-1995

APPLICATION NUMBER: 60/018,496

FILING DATE: 29-MAY-1996

APPLICATION NUMBER: 08/690,011

FILING DATE: 31-JULY-1996

ATTORNEY/AGENT INFORMATION:

NAME: Serunian, Leslie A.

REGISTRATION NUMBER: 35,353

REFERENCE/DOCKET NUMBER: 2026-4199US2

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212)758-4800

TELEFAX: (212)751-6849

INFORMATION FOR SEQ ID NO: 32:

SEQUENCE CHARACTERISTICS:

LENGTH: 296 base pairs

TYPE: nucleic acid

STRANDEDNESS: unknown

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

SEQUENCE DESCRIPTION: SEQ ID NO: 32:

US-09-299-495F-32

Alignment Scores:
Pred. No.: 45.4 Length: 296
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 16.28% Indels: 0
DB: 3 Gaps: 0

US-10-624-218-4 (1-43) x US-09-299-495F-32 (1-296)

Qy 20 AlacluLeuGluGlnGluVal 26
Db 130 GCTGAACCTCGAGCAGGAAGTG 150

RESULT 10

US-09-949-016-50871/c

Sequence 50871, Application US/09949016

Patent No. 6812339

GENERAL INFORMATION:

APPLICANT: VENTER, J. Craig et al.

TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

FILE REFERENCE: CL001307

CURRENT APPLICATION NUMBER: US/09/949,016

CURRENT FILING DATE: 2000-04-14

PRIOR APPLICATION NUMBER: 60/241,755

PRIOR FILING DATE: 2000-10-20

PRIOR APPLICATION NUMBER: 60/237,768

PRIOR FILING DATE: 2000-10-03

PRIOR APPLICATION NUMBER: 60/231,498

PRIOR FILING DATE: 2000-09-08

NUMBER OF SEQ ID NOS: 207012

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 50871

LENGTH: 601

TYPE: DNA

ORGANISM: Human

US-09-949-016-50871

Alignment Scores:
Pred. No.: 88.1 Length: 601
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 16.28% Indels: 0
DB: 4 Gaps: 0

US-10-624-218-4 (1-43) x US-09-949-016-50871 (1-601)

Qy 9 ArgGlnArgAsnThrAlaLeu 15

Db 222 AGCAGAGAAACACTGCACGTG 202

RESULT 11

US-09-949-016-50872/c

Sequence 50872, Application US/09949016

Patent No. 6812339

GENERAL INFORMATION:

APPLICANT: VENTER, J. Craig et al.

TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

FILE REFERENCE: CL001307

CURRENT APPLICATION NUMBER: US/09/949,016

CURRENT FILING DATE: 2000-04-14

PRIOR APPLICATION NUMBER: 60/241,755

PRIOR FILING DATE: 2000-10-20

PRIOR APPLICATION NUMBER: 60/237,768

PRIOR FILING DATE: 2000-10-03

PRIOR APPLICATION NUMBER: 60/231,498

PRIOR FILING DATE: 2000-09-08

NUMBER OF SEQ ID NOS: 207012

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 50872

LENGTH: 601

TYPE: DNA

ORGANISM: Human

US-09-949-016-50872

Alignment Scores:
Pred. No.: 88.1 Length: 601
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0

Query Match: 16.28% Indels: 0
DB: 4 Gaps: 0

US-10-624-218-4 (1-43) x US-09-949-016-50872 (1-601)

Qy 9 ArgGlnArgAenThrAlaLeu 15
Db 104 AGGCAGAGAACTGCACTG 84

RESULT 12

US-09-949-016-93464
; Sequence 93464, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 93464
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-93464

Alignment Scores:
Pred. No.: 89.1 Length: 601
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 16.28% Indels: 0
DB: 4 Gaps: 0

US-10-624-218-4 (1-43) x US-09-949-016-93464 (1-601)

Qy 12 AenThrAlaLeuArgThrGlu 18
Db 536 AATACTGCCCTAAGAAGCTGAG 556

RESULT 13

US-09-479-309-8
; Sequence 8, Application US/09479309
; Patent No. 6110691
; GENERAL INFORMATION:
; APPLICANT: Wang, Xiaodong
; TITLE OF INVENTION: Activators of Caspases
; FILE REFERENCE: UTS0630
; CURRENT APPLICATION NUMBER: US/09/479,309
; CURRENT FILING DATE: 2000-01-06
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 720
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-479-309-8

Alignment Scores:
Pred. No.: 104 Length: 720
Score: 7.00 Matches: 7

Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 16.28% Indels: 0
DB: 3 Gaps: 0

US-10-624-218-4 (1-43) x US-09-479-309-8 (1-720)

Qy 20 AlaGluLeuGluGlnGluVal 26
Db 682 GCTGAGTTAGACGAGGAGTA 702

RESULT 14

US-09-627-393-8
; Sequence 8, Application US/09627393
; Patent No. 6534267
; GENERAL INFORMATION:
; APPLICANT: Wang, Xiaodong
; APPLICANT: Du, Chunying
; TITLE OF INVENTION: Activators of Caspases
; FILE REFERENCE: UTS0630
; CURRENT APPLICATION NUMBER: US/09/627,393
; CURRENT FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: 09/479,309
; PRIOR FILING DATE: 2000-01-06
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 720
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-627-393-8

Alignment Scores:
Pred. No.: 104 Length: 720
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 16.28% Indels: 0
DB: 4 Gaps: 0

US-10-624-218-4 (1-43) x US-09-627-393-8 (1-720)

Qy 20 AlaGluLeuGluGlnGluVal 26
Db 682 GCTGAGTTAGACGAGGAGTA 702

RESULT 15

US-08-428-091-3
; Sequence 3, Application US/08428091
; Patent No. 5683890
; GENERAL INFORMATION:
; APPLICANT: GERMOND, JACQUES.-EDOUARD
; APPLICANT: MARCISSET, OLIVIER
; APPLICANT: MOLLET, BEAT
; TITLE OF INVENTION: BACTERIOCINS OF STREPTOCOCCUS
; THERMOPHILUS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 AVENUE OF THE AMERICA
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:

```

/ APPLICATION NUMBER: US/08/428,091
/ FILING DATE:
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: EP PCT/EP94/02805
/ FILING DATE: 24-AUG-1994
/ APPLICATION NUMBER: CH 2628/93-7
/ FILING DATE: 03-SEP-1993
/ ATTORNEY/AGENT INFORMATION:
/ NAME: FANUCCI, ALLAN A
/ REGISTRATION NUMBER: 30256
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 212 790 9090
/ TELEFAX: 212 869 8864
/ INFORMATION FOR SEQ ID NO: 3:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 770 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA (genomic)
/ ORIGINAL SOURCE:
/ ORGANISM: Streptococcus thermophilus
/ STRAIN: CNCM I-1351
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 221..475
/ FEATURE:
/ NAME/KEY: sig_peptide
/ LOCATION: 221..289
/ FEATURE:
/ NAME/KEY: mat_peptide
/ LOCATION: 290..475
/ OTHER INFORMATION: /function= "encodes for
/ OTHER INFORMATION: thermophiline 1"
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 495..686
/ FEATURE:
/ NAME/KEY: sig_peptide
/ LOCATION: 495..557
/ FEATURE:
/ NAME/KEY: mat_peptide
/ LOCATION: 558..686
/ OTHER INFORMATION: /function= "encodes for
/ OTHER INFORMATION: thermophiline 2"
/ US-08-428-091-3

```

```

Alignment Scores:
Pred. No.: 111 Length: 770
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 16.28% Indels: 0
DB: 1 Gaps: 0

```

US-10-624-218-4 (1-43) x US-08-428-091-3 (1-770)

```

Qy      24  GlnGluValGlnArgLeuGlu 30
Db      622  CAGGAGGTGCAGCGCTGGAG 642

```

Search completed: July 27, 2005, 16:49:39
Job time : 127.333 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 27, 2005, 11:55:09 ; Search time 546.16 Seconds
(without alignments)
509.100 Million cell updates/sec

Title: US-10-624-218-4

Perfect score: 43

Sequence: 1 LETRAAFIRQNTALRTEVA.....QEVORLENEVSQVETRYGPL 43

Scoring table:

OLIGO Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 7277826 seqs, 3233139505 residues

Word size: 1

Total number of hits satisfying chosen parameters: 14542843

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool/p/US10624218/runat_26072005_121437_3743/app_query.fasta_1.796
-DB=Published Applications_NA -Qfmt=fastap -SUFFIX=olip2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=oligo
-TRANS=human40.cdi -LIST=45 -DOCLIGN=200 -THR SCORE=quality -THR MIN=1
-ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10624218 @CGN 1.1189 @runat_26072005_121437_3743
-NCPU=6 -ICPU=3 -NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60
-FPGAPOP=6 -FPGAPEXT=7 -YGAPOP=60 -YGAPEXT=6 -DELOP=6 -DELEXT=7

Database : Published Applications_NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
19: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq.*
20: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq.*
21: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq.*
22: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
23: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq.*
24: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
25: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
26: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	9	20.9	475	10	US-09-918-995-12780 Sequence 12780, A
2	9	20.9	499	20	US-10-357-930-15110 Sequence 15110, A
3	9	20.9	509	20	US-10-357-930-23778 Sequence 23778, A
4	9	20.9	509	20	US-10-357-930-29682 Sequence 29682, A
5	9	20.9	526	20	US-10-357-930-5941 Sequence 5941, Ap
6	9	20.9	617	20	US-10-357-930-35938 Sequence 35938, A
7	9	20.9	617	20	US-10-357-930-45002 Sequence 45002, A
8	9	20.9	1508	18	US-10-092-900A-227 Sequence 227, App
9	9	20.9	1776	18	US-10-092-900A-225 Sequence 225, App
10	9	20.9	1818	17	US-10-094-749-522 Sequence 522, App
11	9	20.9	7233	19	US-10-437-963-96866 Sequence 96866, A
12	8	18.6	25	21	US-10-719-900-752882 Sequence 752882, A
13	8	18.6	25	21	US-10-809-189-96364 Sequence 96364, A
14	8	18.6	542	19	US-10-021-323-15743 Sequence 15743, A
15	8	18.6	557	16	US-10-029-386-7467 Sequence 7467, Ap
16	8	18.6	600	22	US-10-972-079-81913 Sequence 81913, A
17	8	18.6	627	19	US-10-437-963-73527 Sequence 73527, A
18	8	18.6	651	18	US-10-424-599-30989 Sequence 30989, A
19	8	18.6	791	18	US-10-424-599-89140 Sequence 89140, A
20	8	18.6	2280	19	US-10-437-963-73548 Sequence 73548, A
21	7	16.3	25	21	US-10-719-900-771761 Sequence 771761, A
22	7	16.3	25	22	US-10-719-956-326679 Sequence 326679, A
23	7	16.3	168	9	US-09-923-876-395 Sequence 395, App
24	7	16.3	168	10	US-09-923-876-395 Sequence 395, App
25	7	16.3	181	20	US-10-425-115-142296 Sequence 142296, A
26	7	16.3	199	9	US-09-867-701-2609 Sequence 2609, Ap
27	7	16.3	201	20	US-10-719-993-18531 Sequence 18531, A
28	7	16.3	201	20	US-10-719-993-18532 Sequence 18532, A
29	7	16.3	201	20	US-10-719-993-23475 Sequence 23475, A
30	7	16.3	201	20	US-10-719-993-23477 Sequence 23477, A
31	7	16.3	201	21	US-10-741-600-38366 Sequence 38366, A
32	7	16.3	201	21	US-10-741-600-38367 Sequence 38367, A
33	7	16.3	201	21	US-10-741-600-43308 Sequence 43308, A
34	7	16.3	201	21	US-10-741-600-43310 Sequence 43310, A
35	7	16.3	207	18	US-10-333-895-8 Sequence 8, Appli
36	7	16.3	220	17	US-10-182-230-50 Sequence 50, Appl
37	7	16.3	257	9	US-09-867-701-2637 Sequence 2637, Ap
38	7	16.3	296	14	US-10-059-720-32 Sequence 32, Appl
39	7	16.3	304	19	US-10-767-701-30313 Sequence 30313, A
40	7	16.3	329	13	US-10-040-739-806 Sequence 806, App
41	7	16.3	330	9	US-09-880-107-893 Sequence 893, App
42	7	16.3	354	21	US-10-928-992-61 Sequence 61, Appl
43	7	16.3	373	20	US-10-425-115-122218 Sequence 122218, A
44	7	16.3	388	19	US-10-437-963-9094 Sequence 9094, Ap
45	7	16.3	405	20	US-10-425-115-82929 Sequence 82929, A

ALIGNMENTS

RESULT 1
US-09-918-995-12780
; Sequence 12780, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 12780

```
; LENGTH: 475
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(475)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-12780

Alignment Scores:
Pred. No.: 2.37 Length: 475
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 20.93% Indels: 0
DB: 10 Gaps: 0

US-10-624-218-4 (1-43) x US-09-918-995-12780 (1-475)

Qy 21 GluLeuGluGlnGluValGlnArgLeu 29
|||
Db 160 GAACTGGAGCAAGAGGTGCAAGACTT 186

RESULT 2
US-10-357-930-15110
; Sequence 15110, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15110
; LENGTH: 499
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 508, 509
; OTHER INFORMATION: n = A,T,C or G
US-10-357-930-23778

Alignment Scores:
Pred. No.: 2.52 Length: 509
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 20.93% Indels: 0
DB: 20 Gaps: 0

US-10-624-218-4 (1-43) x US-10-357-930-23778 (1-509)

Qy 21 GluLeuGluGlnGluValGlnArgLeu 29
|||
Db 5 GAACTGGAGCAAGAGGTGCAAGACTT 31

RESULT 4
US-10-357-930-29682
; Sequence 29682, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
```


;; PRIOR FILING DATE: 2000-03-16
;; PRIOR APPLICATION NUMBER: 60/207,454
;; PRIOR FILING DATE: 2000-05-25
;; PRIOR APPLICATION NUMBER: 60/211,314
;; PRIOR FILING DATE: 2000-06-09
;; PRIOR APPLICATION NUMBER: 60/219,007
;; PRIOR FILING DATE: 2000-07-18
;; PRIOR APPLICATION NUMBER: 60/255,281
;; PRIOR FILING DATE: 2000-12-13
;; NUMBER OF SEQ ID NOS: 62232
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 29682
;; LENGTH: 509
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: 508..509
;; OTHER INFORMATION: n = A,T,C or G
US-10-357-930-29682

Alignment Scores:
Pred. No.: 2.52 Length: 509
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 20.93% Indels: 0
DB: 20 Gaps: 0

US-10-624-218-4 (1-43) x US-10-357-930-29682 (1-509)

Qy 21 GluleuGlInGluValGlnArgIeu 29
|||||

Db 5 GAACGGAGCAAGAGGTGCAAGACTT 31

RESULT 5

US-10-357-930-5941
;; Sequence 5941, Application US/10357930
;; Publication No. US20040259086A1
;; GENERAL INFORMATION:
;; APPLICANT: Schlegel, Robert
;; APPLICANT: Endege, Wilson
;; APPLICANT: Monahan, John
;; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
;; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
;; TITLE OF INVENTION: HUMAN PROSTATE CANCER

;; FILE REFERENCE: MRI-007BCN
;; CURRENT APPLICATION NUMBER: US/10/357,930
;; CURRENT FILING DATE: 2003-02-04
;; PRIOR APPLICATION NUMBER: 09/785,276
;; PRIOR FILING DATE: 2003-02-16
;; PRIOR APPLICATION NUMBER: 60/183,319
;; PRIOR FILING DATE: 2000-02-17
;; PRIOR APPLICATION NUMBER: 60/189,862
;; PRIOR FILING DATE: 2000-03-16
;; PRIOR APPLICATION NUMBER: 60/207,454
;; PRIOR FILING DATE: 2000-05-25
;; PRIOR APPLICATION NUMBER: 60/211,314
;; PRIOR FILING DATE: 2000-06-09
;; PRIOR APPLICATION NUMBER: 60/219,007
;; PRIOR FILING DATE: 2000-07-18
;; PRIOR APPLICATION NUMBER: 60/255,281
;; PRIOR FILING DATE: 2000-12-13
;; NUMBER OF SEQ ID NOS: 62232
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 5941
;; LENGTH: 526
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-10-357-930-5941

Alignment Scores:
Pred. No.: 2.59 Length: 526

Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 20.93% Indels: 0
DB: 20 Gaps: 0

US-10-624-218-4 (1-43) x US-10-357-930-5941 (1-526)

Qy 21 GluleuGlInGluValGlnArgIeu 29
|||||

Db 32 GAACGGAGCAAGAGGTGCAAGACTT 58

RESULT 6

US-10-357-930-35938
;; Sequence 35938, Application US/10357930
;; Publication No. US20040259086A1
;; GENERAL INFORMATION:
;; APPLICANT: Schlegel, Robert
;; APPLICANT: Endege, Wilson
;; APPLICANT: Monahan, John
;; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
;; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
;; TITLE OF INVENTION: HUMAN PROSTATE CANCER

;; FILE REFERENCE: MRI-007BCN
;; CURRENT APPLICATION NUMBER: US/10/357,930
;; CURRENT FILING DATE: 2003-02-04
;; PRIOR APPLICATION NUMBER: 09/785,276
;; PRIOR FILING DATE: 2003-02-16
;; PRIOR APPLICATION NUMBER: 60/183,319
;; PRIOR FILING DATE: 2000-02-17
;; PRIOR APPLICATION NUMBER: 60/189,862
;; PRIOR FILING DATE: 2000-03-16
;; PRIOR APPLICATION NUMBER: 60/207,454
;; PRIOR FILING DATE: 2000-05-25
;; PRIOR APPLICATION NUMBER: 60/211,314
;; PRIOR FILING DATE: 2000-06-09
;; PRIOR APPLICATION NUMBER: 60/219,007
;; PRIOR FILING DATE: 2000-07-18
;; PRIOR APPLICATION NUMBER: 60/255,281
;; PRIOR FILING DATE: 2000-12-13
;; NUMBER OF SEQ ID NOS: 62232
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 35938
;; LENGTH: 617
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-10-357-930-35938

Alignment Scores:
Pred. No.: 2.97 Length: 617
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 20.93% Indels: 0
DB: 20 Gaps: 0

US-10-624-218-4 (1-43) x US-10-357-930-35938 (1-617)

Qy 21 GluleuGlInGluValGlnArgIeu 29
|||||

Db 113 GAACGGAGCAAGAGGTGCAAGACTT 139

RESULT 7

US-10-357-930-45002
;; Sequence 45002, Application US/10357930
;; Publication No. US20040259086A1
;; GENERAL INFORMATION:
;; APPLICANT: Schlegel, Robert
;; APPLICANT: Endege, Wilson
;; APPLICANT: Monahan, John
;; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
;; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
;; TITLE OF INVENTION: HUMAN PROSTATE CANCER

; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 45002
; LENGTH: 617
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-357-930-45002

Alignment Scores:
Pred. No.: 2,97 Length: 617
Score: 9,00 Matches: 9
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 20.93% Indels: 0
DB: 20 Gaps: 0

US-10-624-218-4 (1-43) x US-10-357-930-45002 (1-617)

Qy 21 GluLeuGluGlnGluValGlnArgLeu 29
Db 113 GAACGGAGCAAGAGGTGCAAGACTT 139

RESULT 8
US-10-092-900A-227
; Sequence 227, Application US/10092900A
; Publication No. US20040043382A1
; GENERAL INFORMATION:
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Taupier Jr., Raymond J.
; APPLICANT: Pena, Carol E.A.
; APPLICANT: Li, Li
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Ji, Weizhen
; APPLICANT: Gorman, Linda
; APPLICANT: Miller, Charles E.
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Patturajan, Meera
; APPLICANT: Gangolli, Esha A.M.
; APPLICANT: Vernet, Corine A.M.
; APPLICANT: Tchernev, Velizar T.
; APPLICANT: Fernandes, Elma R.
; APPLICANT: Casman, Stacie J.
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Liu, Yi
; APPLICANT: Anderson, David W.
; APPLICANT: Spaderna, Steven K.
; APPLICANT: Catterton, Elina
; APPLICANT: Leite, Mario W.
; APPLICANT: Zhong, Haihong
; APPLICANT: Alsobrook, John P.
; APPLICANT: Lepley, Denise M.

; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: No. US20040043382A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-290C
; CURRENT APPLICATION NUMBER: US/10/092,900A
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: USSN 60/274,322
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: USSN 60/283,675
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: USSN 60/338,092
; PRIOR FILING DATE: 2001-12-03
; PRIOR APPLICATION NUMBER: USSN 60/274,281
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: USSN 60/274,191
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: USSN 60/325,681
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: USSN 60/304,354
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: USSN 60/279,995
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: USSN 60/294,899
; PRIOR FILING DATE: 2001-05-31
; PRIOR APPLICATION NUMBER: USSN 60/287,424
; PRIOR FILING DATE: 2001-04-30
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 768
; SEQ ID NO 227
; LENGTH: 1508
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2)..(1445)
US-10-092-900A-227

Alignment Scores:
Pred. No.: 6,45 Length: 1508
Score: 9,00 Matches: 9
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 20.93% Indels: 0
DB: 18 Gaps: 0

US-10-624-218-4 (1-43) x US-10-092-900A-227 (1-1508)

Qy 21 GluLeuGluGlnGluValGlnArgLeu 29
Db 59 GAACGGAGCAAGAGGTGCAAGACTT 85

RESULT 9
US-10-092-900A-225
; Sequence 225, Application US/10092900A
; Publication No. US20040043382A1
; GENERAL INFORMATION:
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Taupier Jr., Raymond J.
; APPLICANT: Pena, Carol E.A.
; APPLICANT: Li, Li
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Ji, Weizhen
; APPLICANT: Gorman, Linda
; APPLICANT: Miller, Charles E.
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Patturajan, Meera
; APPLICANT: Gangolli, Esha A.
; APPLICANT: Vernet, Corine A.M.
; APPLICANT: Guo, Xiaojia Sasha
; APPLICANT: Tchernev, Velizar T.

APPLICANT: Fernandes, Elma R.
APPLICANT: Casman, Stacie J.
APPLICANT: Malyankar, Uriel M.
APPLICANT: Gerlach, Valerie
APPLICANT: Liu, Yi
APPLICANT: Anderson, David W.
APPLICANT: Spaderna, Steven K.
APPLICANT: Catterton, Elina
APPLICANT: Leite, Mario W.
APPLICANT: Zhong, Haihong
APPLICANT: Alsobrook, John P.
APPLICANT: Lepley, Denise M.
APPLICANT: Rieger, Daniel K.
APPLICANT: Burgess, Catherine E.
TITLE OF INVENTION: NO. US20040043382A1el Proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 21402-290C
CURRENT APPLICATION NUMBER: US/10/092,900A
CURRENT FILING DATE: 2002-03-07
PRIOR APPLICATION NUMBER: USSN 60/274,322
PRIOR FILING DATE: 2001-03-08
PRIOR APPLICATION NUMBER: USSN 60/283,675
PRIOR FILING DATE: 2001-04-13
PRIOR APPLICATION NUMBER: USSN 60/338,092
PRIOR FILING DATE: 2001-12-03
PRIOR APPLICATION NUMBER: USSN 60/274,281
PRIOR FILING DATE: 2001-03-08
PRIOR APPLICATION NUMBER: USSN 60/274,191
PRIOR FILING DATE: 2001-03-08
PRIOR APPLICATION NUMBER: USSN 60/325,681
PRIOR FILING DATE: 2001-09-27
PRIOR APPLICATION NUMBER: USSN 60/304,354
PRIOR FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: USSN 60/279,995
PRIOR FILING DATE: 2001-03-30
PRIOR APPLICATION NUMBER: USSN 60/294,899
PRIOR FILING DATE: 2001-05-31
PRIOR APPLICATION NUMBER: USSN 60/287,424
PRIOR FILING DATE: 2001-04-30
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 768
SEQ ID NO 225
LENGTH: 1776
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (31)..(1474)
US-10-092-900A-225

Alignment Scores:
Pred. No.: 7.43 Length: 1776
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 20.93% Indels: 0
DB: 18 Gaps: 0

US-10-624-218-4 (1-43) x US-10-092-900A-225 (1-1776)
Qy 21 GluLeuGluGlnValGlnArgLeu 29
Db 88 GAACGGAGCAAGAGGTGCAAGACTT 114

RESULT 10
US-10-094-749-522
Sequence 522, Application US/10094749
Publication No. US20030219741A1
GENERAL INFORMATION:
APPLICANT: ISOGAI, TAKAO
APPLICANT: SUGIYAMA, TOMOYASU
APPLICANT: OTSUKI, TETSUJI
APPLICANT: WAKAMATSU, AI
APPLICANT: SATO, HIROYUKI

APPLICANT: ISHII, SHIZUKO
APPLICANT: YAMAMOTO, JUN-ICHI
APPLICANT: ISONO, YUUKO
APPLICANT: HIO, YURI
APPLICANT: OTSUKA, KAORU
APPLICANT: NAGAI, KEIICHI
APPLICANT: IRIE, RYOTARO
APPLICANT: TAMECHIKA, ICHIRO
APPLICANT: SEKI, NAOHICO
APPLICANT: YOSHIKAWA, TSUTOMU
APPLICANT: OTSUKA, MOTOUYUKI
APPLICANT: NAGAHARI, KENJI
APPLICANT: MASUHO, YASUHIKO
TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
FILE REFERENCE: 084335/0160
CURRENT APPLICATION NUMBER: US/10/094,749
CURRENT FILING DATE: 2002-03-12
PRIOR APPLICATION NUMBER: 60/350,435
PRIOR FILING DATE: 2002-01-24
PRIOR APPLICATION NUMBER: JP 2001-328381
PRIOR FILING DATE: 2001-09-14
NUMBER OF SEQ ID NOS: 3381
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 522
LENGTH: 1818
TYPE: DNA
ORGANISM: Homo sapiens
US-10-094-749-522

Alignment Scores:
Pred. No.: 7.58 Length: 1818
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 20.93% Indels: 0
DB: 17 Gaps: 0

US-10-624-218-4 (1-43) x US-10-094-749-522 (1-1818)
Qy 21 GluLeuGluGlnValGlnArgLeu 29
Db 133 GAACGGAGCAAGAGGTGCAAGACTT 159

RESULT 11
US-10-437-963-96866/c
Sequence 96866, Application US/10437963
Publication No. US20040123343A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Wu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbazuk, Brad
APPLICANT: Li, Ping
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 96866
LENGTH: 7233
TYPE: DNA
ORGANISM: Oryza sativa
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_94922C.1
US-10-437-963-96866

Alignment Scores:
Pred. No.: 25.1 Length: 7233
Score: 9.00 Matches: 9

```
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%    Mismatches: 0
Query Match: 20.93%              Indels: 0
DB: 19                           Gaps: 0

US-10-624-218-4 (1-43) x US-10-437-963-96866 (1-7233)

QY 18 GluValAlaGluLeuGluGlnGluVal 26
   |||||
Db 353 GAGGTGGCGAGCTGGAGCAGAGGTG 327

RESULT 12
US-10-719-900-752882/c
; Sequence 752882, Application US/10719900
; Publication No. US2005002616A1
; GENERAL INFORMATION:
; APPLICANT: Xue Mei Zhou
; TITLE OF INVENTION: Methods of Genetic Analysis of Mouse
; FILE REFERENCE: 3528.1
; CURRENT APPLICATION NUMBER: US/10/719,900
; CURRENT FILING DATE: 2003-11-20
; PRIOR APPLICATION NUMBER: 60/427,808
; PRIOR FILING DATE: 2002-11-20
; NUMBER OF SEQ ID NOS: 982914
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 752882
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-719-900-752882

Alignment Scores:
Pred. No.: 2.21      Length: 25
Score: 8.00          Matches: 8
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%    Mismatches: 0
Query Match: 18.60%              Indels: 0
DB: 21               Gaps: 0

US-10-624-218-4 (1-43) x US-10-719-900-752882 (1-25)

QY 1 LeuGluIleArgAlaAlaPheLeu 8
   |||||
Db 25 CTGGAGATCAGAGCAGCCTTCCTT 2

RESULT 13
US-10-809-189-96364
; Sequence 96364, Application US/10809189
; Publication No. US20050048531A1
; GENERAL INFORMATION:
; APPLICANT: Michael Mittmann
; APPLICANT: David Mack
; APPLICANT: David Lockhart
; APPLICANT: Affymetrix, Inc.
; TITLE OF INVENTION: Methods of Genetic Analysis
; FILE REFERENCE: 3101.1
; CURRENT APPLICATION NUMBER: US/10/809,189
; CURRENT FILING DATE: 2004-03-25
; PRIOR APPLICATION NUMBER: US/09/396,196
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: 60/100,678
; PRIOR FILING DATE: 1998-09-17
; NUMBER OF SEQ ID NOS: 127806
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 96364
; LENGTH: 25
; TYPE: DNA
; ORGANISM: mus musculus
US-10-809-189-96364

Alignment Scores:
Pred. No.: 2.21      Length: 25
Score: 8.00          Matches: 8
```

```
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%    Mismatches: 0
Query Match: 18.60%              Indels: 0
DB: 21                           Gaps: 0

US-10-624-218-4 (1-43) x US-10-809-189-96364 (1-25)

QY 15 LeuArgThrGluValAlaGluLeu 22
   |||||
Db 1 CTGCGGACGGAGGTTCGCGAGCTT 24

RESULT 14
US-10-021-323-15743/c
; Sequence 15743, Application US/10021323
; Publication No. US20040123340A1
; GENERAL INFORMATION:
; APPLICANT: Deikman, Jill
; APPLICANT: Feng, Paul C.C.
; APPLICANT: Fincher, Karen L.
; APPLICANT: Ziegler, Todd E.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(52274)B
; CURRENT APPLICATION NUMBER: US/10/021,323
; CURRENT FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: US 60/255, 619
; PRIOR FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 17880
; SEQ ID NO 15743
; LENGTH: 542
; TYPE: DNA
; ORGANISM: Gossypium hirsutum
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3829-030-Q6-K6-B12
US-10-021-323-15743

Alignment Scores:
Pred. No.: 31.7      Length: 542
Score: 8.00          Matches: 8
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%    Mismatches: 0
Query Match: 18.60%              Indels: 0
DB: 19               Gaps: 0

US-10-624-218-4 (1-43) x US-10-021-323-15743 (1-542)

QY 11 ArgAsnThrAlaLeuArgThrGlu 18
   |||||
Db 312 CGTAACACTGCACCTCGAACTGAG 289

RESULT 15
US-10-029-386-7467
; Sequence 7467, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR GI
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 7467
; LENGTH: 557
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR18.3
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.5
```

; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2
; OTHER INFORMATION: NT HIT: q17656917, EVALUE 1.00e-35
; OTHER INFORMATION: EST HUMAN HIT: A1335888.1, EVALUE 2.00e-87
; OTHER INFORMATION: SWISSPROT HIT: P98195, EVALUE 2.00e-21
US-10-029-386-7467

Alignment Scores:
Pred. No.: 32.5 Length: 557
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 18.60% Indels: 0
DB: 16 Gaps: 0

US-10-624-218-4 (1-43) x US-10-029-386-7467 (1-557)

Qy 4 ArgAlaIaPheLeuArgGlnArg 11
Db 298 CGTGGGCTTTCCTGAGGCGG 321

Search completed: July 27, 2005, 17:17:54
Job time : 550.16 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 27, 2005, 11:14:04 ; Search time 2989.69 Seconds
(without alignments)
547.469 Million cell updates/sec

Title: US-10-624-218-4

Perfect score: 43

Sequence: 1 LEIRAAFLQRNTALRTEVA.....QEVORLENSQVETRYGPL 43

Scoring table:

```
OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0
```

Searched: 34239544 seqs, 19032134700 residues

Word size: 1

Total number of hits satisfying chosen parameters: 68475752

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

```
-MODEL=frame+ p2n.model -DEV=xlp
-Q/cgn2_1/USPRO_spool_p/US10624218/runat_26072005_121436_3698/app_query.fasta_1.796
-DB=EST -QFMT=fastcap -SUFFIX=olip2n.rst -MINWATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=oligo -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10624218.ecgn.1.1.9525 @runat_26072005_121436_3698 -NCPU=6 -ICPU=3
-NO_MMALP -LARGOQUERY -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7
```

Database :

```
EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gsi1:*
9: gb_gsi2:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	11	25.6	341	1 A1014348	A1014348 am51g11.s
2	11	25.6	351	6 BY775604	BY775604 BY775604
3	11	25.6	380	1 A1892971	A1892971 mq82h05.Y
4	11	25.6	398	6 CB707074	CB707074 AMGNNUC.T
5	11	25.6	418	7 CN394275	CN394275 170004553
6	11	25.6	440	2 BE295612	BE295612 601176577
7	11	25.6	446	1 A1138848	A1138848 mq82h05.r
8	11	25.6	464	4 BG986696	BG986696 PM1-HT117
9	11	25.6	484	5 BU262464	BU262464 603502757

C	10	11	25.6	498	2 AM836776	AM836776 QV1-LT003
	11	11	25.6	528	2 BF660216	BF660216 maa29a04.
	12	11	25.6	540	7 CR536622	CR536622 DKF2p459A
C	13	11	25.6	555	2 AM836791	AM836791 QV1-LT003
	14	11	25.6	557	5 BU311019	BU311019 603540871
	15	11	25.6	563	7 CO794673	CO794673 NT144B.B0
	16	11	25.6	578	1 AL120537	AL120537 DKF2p761A
	17	11	25.6	599	2 AW500333	AW500333 UI-HF-BNO
	18	11	25.6	619	2 BF306473	BF306473 601893367
	19	11	25.6	626	5 BU469699	BU469699 603761076
	20	11	25.6	631	6 CB577428	CB577428 AMGNNUC.N
	21	11	25.6	638	7 CK624751	CK624751 mi24g09.Y
	22	11	25.6	653	2 BF203280	BF203280 601866155
	23	11	25.6	654	4 BG432348	BG432348 602496669
C	24	11	25.6	658	9 CR828685	CR828685 GROAA60A
	25	11	25.6	673	5 BM963436	BM963436 UI-M-EQO-
	26	11	25.6	693	6 CB247229	CB247229 UI-M-F10-
	27	11	25.6	786	5 BU466941	BU466941 603371693
	28	11	25.6	788	7 CN455515	CN455515 UI-M-HNO-
	29	11	25.6	812	5 BQ042190	BQ042190 UI-M-EQO-
	30	11	25.6	850	2 BF309490	BF309490 601892191
	31	11	25.6	928	6 CB590113	CB590113 AGENCOURT
	32	11	25.6	933	5 BQ887766	BQ887766 AGENCOURT
	33	11	25.6	941	2 BE737088	BE737088 601304854
	34	11	25.6	958	5 BU149688	BU149688 AGENCOURT
	35	9	20.9	162	1 AA376153	AA376153 EST88536
	36	9	20.9	462	1 AV747602	AV747602 AV747602
	37	9	20.9	509	4 BM759156	BM759156 K-EST0039
	38	9	20.9	528	4 BM845153	BM845153 K-EST0123
	39	9	20.9	570	5 BP212591	BP212591 BP212591
	40	9	20.9	581	5 BP224795	BP224795 BP224795
	41	9	20.9	582	5 BP199893	BP199893 BP199893
	42	9	20.9	582	5 BP200065	BP200065 BP200065
	43	9	20.9	583	5 BP263968	BP263968 BP263968
	44	9	20.9	601	6 CB152685	CB152685 K-EST0209
	45	9	20.9	605	4 BM786445	BM786445 K-EST0065

ALIGNMENTS

RESULT 1

A1014348

LOCUS

DEFINITION

IMAGE:1539140 3', similar to gb:M95586 TRANSCRIPTION FACTOR E2-ALPHA

(HUMAN); mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

A1014348 341 bp mRNA linear EST 15-JUN-1998
am51g11.b1 Johnston frontal cortex Homo sapiens cDNA clone
IMAGE:1539140 3', similar to gb:M95586 TRANSCRIPTION FACTOR E2-ALPHA
(HUMAN); mRNA sequence.
A1014348 GI:3228180
A1014348 EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 341)
Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S.,
Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M.,
Martin, J., Moore, B., Schellenger, K., Steptoe, M., Tan, F.,
Theising, B., White, Y., Wyllie, T., Waterston, R. and Wilson, R.
WashU-NCI human EST Project
Unpublished (1997)
Contact: Wilton RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 314.
Location/Qualifiers
1. 341
/organism="Homo sapiens"

```

/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1539140"
/sex="male"
/tissue_type="pooled frontal lobe"
/dev_stage="adult"
/lab_host="SOLR (kanamycin resistant)"
/clone_lib="Johnston frontal cortex"
/notes="Organ: brain; Vector: Bluescript SK-; Site_1: EcORI; Stanley Neuropathology Consortium (www.stanleylab.org) brains S-58, S-65, S-67, S-78. Random + oligo-dT primed into EcORI site of ZAP II Vector. Mass excised. Avg insert length 1.9kb. Custom library provided by Dr. Nancy Johnston [(410) 614-3918, nlj@welchlink.welch.jhu.edu]."
```

ORIGIN

Alignment Scores:

Pred. No.:	0.0805	Length:	341
Score:	11.00	Matches:	11
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	25.58%	Indels:	0
DB:	1	Gaps:	0

US-10-624-218-4 (1-43) x A1014348 (1-341)

QY 12 AsnThrAlaLeuArgThrGluValAlaGluLeu 22
|||||

Db 155 AACACAGCCCTGCGCAGGAGGTGGCGGAGCTA 187
|||||

RESULT 2

LOCUS BY775604

DEFINITION BY775604 RIKEN full-length enriched, 17.5 days embryo whole body

ACCESSION BY775604.1 GI:39702242

VERSION

KEYWORDS Mus musculus (house mouse)

SOURCE

ORGANISM

REFERENCE

AUTHORS

1 (bases 1 to 351)

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

Carninci, P., Waki, K., Shiraki, T., Konno, H., Shibata, K., Itoh, M., Azawa, K., Arakawa, T., Ishii, Y., Sasaki, D., Bono, H., Kondo, S., Sugahara, Y., Saito, R., Osato, N., Fukuda, S., Sato, K., Wataniki, A., Hirozane-Kishikawa, T., Nakamura, M., Shibata, Y., Yasunishi, A., Kikuchi, N., Yoshiki, A., Kusakabe, M., Gustincich, S., Beisel, K., Pavan, W., Aidinis, V., Nakagawara, A., Held, W. A., Iwata, H., Kono, T., Nakauchi, H., Lyons, P., Wells, C., Hume, D. A., Fagioli, M., Hensch, T. K., Brinkmeier, M., Camper, S., Hirota, J., Mombaerts, P., Muramatsu, M., Okazaki, Y., Kawai, J. and Hayashizaki, Y.

Targeting a complex transcriptome: the construction of the mouse full-length cDNA encyclopedia

Genome Res. 13 (6B), 1273-1289 (2003)

22703353

12819125

CONTACT: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site (<http://genome.gsc.riken.jp/>) for further details.

FEATURES

source

1. .351

/organism="Mus musculus"

/mol_type="mRNA"

/strain="CS7BL/6J"

/db_xref="taxon:10090"

/clone="L930096D17"

/tissue_type="whole body"

/dev_stage="17.5 days embryo"

/clone_lib="RIKEN full-length enriched, 17.5 days embryo whole body"

ORIGIN

Alignment Scores:

Pred. No.:	0.0827	Length:	351
Score:	11.00	Matches:	11
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	25.58%	Indels:	0
DB:	6	Gaps:	0

US-10-624-218-4 (1-43) x BY775604 (1-351)

QY 12 AsnThrAlaLeuArgThrGluValAlaGluLeu 22
|||||

Db 48 AACACAGCCCTGCGCAGGAGGTGGCGGAGCTT 80
|||||

RESULT 3

LOCUS AI892971

DEFINITION AI892971 380 bp mRNA linear EST 15-MAR-2000

ACCESSION AI892971

VERSION

KEYWORDS Mus musculus (house mouse)

SOURCE

ORGANISM

REFERENCE

AUTHORS

1 (bases 1 to 380)

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.

The WashU-NCI Mouse EST Project 1999

Unpublished (1999)

Contact: Marra M/WashU-NCI Mouse EST Project 1999

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: mouseest@watson.wustl.edu

This clone is available royalty-free through LML; contact the IMAGE Consortium (info@image.llnl.gov) for further information.

This read is a RESEQUENCE of a previously sequenced mouse clone

This read has been verified (found to hit its original self in the correct orientation)

MG1:359921

Seq primer: -40RP from Gibco

High quality sequence stop: 336.

FEATURES

source

1. .380

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

/clone="IMAGE:585273"

/tissue_type="melanoma"

/dev_stage="M2 cells"

/lab_host="SOLR (kanamycin resistant)"

/clone_lib="Stratagene mouse melanoma (#937312)"

/notes="Organ: Skin; Vector: pBluescript SK-; Site_1:

EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer:
Oligo dt. From M2 cells, a highly metastatic derivative of
the K-1735 (mouse) melanoma. Average insert size: 1.0 kb;
Uni-ZAP XR Vector; ~5' adaptor sequence: 5' GAATTCGACGAG
3' ~3' adaptor sequence: 5' CTCAGTTTTTTTTTTTTTTT 3'

ORIGIN

Alignment Scores:
Pred. No.: 0.089 Length: 380
Score: 11.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 25.58% Indels: 0
DB: 1 Gaps: 0

US-10-624-218-4 (1-43) x AI892971 (1-380)

Qy 12 AenThrAlaLeuArgThrGluValAlaGluLeu 22
Db 113 AACACAGCCCTGCGACGAGGTTGCTGAGCTT 145

RESULT 4
LOCUS CB707074 398 bp mRNA linear EST 10-APR-2003
DEFINITION AMGNNUC:TRCP2-00001-D6-A trcp2 (10289) Rattus norvegicus cDNA clone
trcp2-00001-d6 5', mRNA sequence.
ACCESSION CB707074
VERSION CB707074.1 GI:29764222
KEYWORDS EST.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus

REFERENCE 1 (bases 1 to 398)
AUTHORS Amgen EST Program.
TITLE Amgen Rat EST Program
JOURNAL Unpublished (2003)
COMMENT Contact: Dan Fitzpatrick
Amgen, Inc
One Amgen Center Drive, Thousand Oaks, CA 91320-1799, USA
Tel: 805 447-4881
Plate: 00001 row: d column: 6.
Location/Qualifiers
1. 398
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/clone="trcp2-00001-d6"
/tissue type="choroid plexus brain"
/clone_lib="trcp2 (10289)"
/note="Vector: C6XFG7L; Site_1: SalI; Site_2: NotI;
choroid plexus brain region"

FEATURES

source

ORIGIN

Alignment Scores:
Pred. No.: 0.093 Length: 398
Score: 11.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 25.58% Indels: 0
DB: 6 Gaps: 0

US-10-624-218-4 (1-43) x CB707074 (1-398)

Qy 12 AenThrAlaLeuArgThrGluValAlaGluLeu 22
Db 74 AACACAGCCCTGCGACGAGGTTGCTGAGCTT 106

RESULT 5
LOCUS CN394275 418 bp mRNA linear EST 16-MAY-2004
DEFINITION 17000455365636 GRN_ES Homo sapiens cDNA 5', mRNA sequence.

ACCESSION CN394275
VERSION CN394275.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 418)
AUTHORS Brandenberger, R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G.J., Li, Y., Xu, C., Fang, R., Guegler, K., Rao, M.S., Mandalam, R., Lebkowski, J. and Stanton, L.W.
TITLE Transcriptome characterization elucidates signaling networks that control human ES cell growth and differentiation
JOURNAL Nat. Biotechnol. 22 (6), 707-716 (2004)
COMMENT Contact: Brandenberger R
Regenerative Medicine
Geron Corporation
230 Constitution Drive, Menlo Park, CA 94025, USA
Tel: 650 473 8658
Fax: 650 473 7760
Email: rbrandenberger@geron.com
Insert Length: 418 Std Error: 0.00.
Location/Qualifiers
1. 418
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="embryonic stem cells, cell lines H1, H7, and H9"
/clone_lib="GRN ES"
/note="Oligo dT primed, full-length enriched cDNA library from undifferentiated hES cell lines H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free conditions"

FEATURES

source

1. 418

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/tissue_type="embryonic stem cells, cell lines H1, H7, and H9"

/clone_lib="GRN ES"

/note="Oligo dT primed, full-length enriched cDNA library from undifferentiated hES cell lines H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free conditions"

Alignment Scores:

Pred. No.: 0.0973 Length: 418

Score: 11.00 Matches: 11

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 25.58% Indels: 0

DB: 7 Gaps: 0

US-10-624-218-4 (1-43) x CN394275 (1-418)

Qy 12 AenThrAlaLeuArgThrGluValAlaGluLeu 22

Db 92 AACACAGCCCTGCGACGAGGTTGCTGAGCTA 124

RESULT 6

BE295612

LOCUS BE295612

DEFINITION 601176577F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:3531762 5', mRNA sequence.

ACCESSION BE295612

VERSION BE295612.1

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 440)

AUTHORS NIH-MGC http://mgc.nci.nih.gov/.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov

Plate: LCM206 row: j column: 19
High quality sequence start: 6
High quality sequence stop: 437.

FEATURES

source

Location/Qualifiers
1. .440
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:585273"
/tissue_type="rhabdomyosarcoma"
/lab_hosts="DHI0B (phage-resistant)"
/clone_lib="NIH_MGC_17"
/note="Organ: muscle; Vector: pOTE7; Site 1: EcoRI;
Site 2: XhoI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."

ORIGIN

Alignment Scores:
Pred. No.: 0.102 Length: 440
Score: 11.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 25.58% Indels: 0
DB: 2 Gaps: 0

US-10-624-218-4 (1-43) x BE295612 (1-440)

Qy 12 AsnThrAlaLeuArgThrGluValAlaGluLeu 22
Db 351 AACACAGCCCTGCGGACGGAGGTGCGGAGCTA 383

RESULT 7

AA138848

LOCUS

DEFINITION AA138848 446 bp mRNA linear EST 09-FEB-1997
mq82805.r1 Stratagene mouse melanoma (#937312) Mus musculus cDNA
Clone IMAGE:585273 5', similar to gb:M95586 TRANSCRIPTION FACTOR
E2-ALPHA (HUMAN); mRNA sequence.

ACCESSION

AA138848

VERSION AA138848.1 GI:1701049

KEYWORDS

EST.

SOURCE Mus musculus (house mouse)

ORGANISM

REFERENCE 1 (bases 1 to 446)
Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
Theising, B., Wylie, T., Lennon, G., Soares, B., Willson, R. and
Waterston, R.

TITLE

JOURNAL

COMMENT

Unpublished (1996)
Contact: Marra M/Mouse EST Project
WashU-HMMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810

Email: mouse@watson.wustl.edu

This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.

MG1:359921

Seq primer: -28m13 rev1 ET from Amersham

High quality sequence stop: 443.

FEATURES

source

Location/Qualifiers
1. .446
/organism="Mus musculus"
/mol_type="mRNA"

/db_xref="taxon:10090"
/clone="IMAGE:585273"
/tissue_type="melanoma"
/dev_stage="M2 cells"
/lab_host="SOLR (kanamycin resistant)"
/clone_lib="Stratagene mouse melanoma (#937312)"
/note="Organ: skin; Vector: pBluescript SK-; Site 1:
EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer:
Oligo dT. From M2 cells, a highly metastatic derivative of
the K-1735 (mouse) melanoma. Average insert size: 1.0 kb;
Uni-ZAP XR Vector; ~5' adaptor sequence: 5' GAATTCGGACGAG
3' ~3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3' "

ORIGIN

Alignment Scores:
Pred. No.: 0.103 Length: 446
Score: 11.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 25.58% Indels: 0
DB: 1 Gaps: 0

US-10-624-218-4 (1-43) x AA138848 (1-446)

Qy 12 AsnThrAlaLeuArgThrGluValAlaGluLeu 22
Db 112 AACACAGCCCTGCGGACGGAGGTGCGGAGCTT 144

RESULT 8

BG986696

LOCUS

DEFINITION BG986696 464 bp mRNA linear EST 13-JUN-2001
PM1-HT1170-030101-001-b09 HT1170 Homo sapiens cDNA, mRNA sequence.

ACCESSION

BG986696

VERSION BG986696.1 GI:14390766

KEYWORDS

EST.

SOURCE Homo sapiens (human)

ORGANISM

REFERENCE

AUTHORS

1 (bases 1 to 464)

Dias, Neco, E., Garcia Correa, R., Verjowski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V.,
O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags

PROC. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE

20202663

PUBMED

10737800

COMMENT

Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tbl=PM1&t2=PM1-HT1170-
030101-001-b09&t3=2001-01-03&t4=1)

Seq primer: puc 18 forward

High quality sequence stop: 6

High quality sequence stop: 396.

FEATURES

source

1. .464

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/dev_stage="Adult"

/clone_lib="HT1170"

(1996): 791, except that a significantly longer reannealing hybridization was used."

ORIGIN

Alignment Scores:			
Pred. No.:	0.112	Length:	484
Score:	11.00	Matches:	11
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	25.58%	Indels:	0
DB:	5	Gaps:	0

U.S.-10-624-218-4 (1-43) x BU262464 (1-484)

Qy 12 AsnThrAlaLeuArgThrGluValAlaGluLeu 22
Db 285 AATACGGCCCTGAGGACGGAGTTGCAGAGCTG 31

RESULT 10
AW836776/c

LOCUS AW836776
DEFINITION QV1-LT0036-150200-070-4
ACCESSION AW836776
VERSION AW836776.1 GI:7930750
KEYWORDS EST.
SOURCE Homo sapiens (human)

RECORDS	SOURCE	ORGANISM	LOC.
		Homo sapiens	(human)
		Homo sapiens	

REFERENCE

AUTHORS
Dias Neto, E., Garcia C
Nagai, M.A., da Silva, W
Cajiao, C. W. Camalho

Goldman, G.H., Carvalho
Brunstein, A., deOliveira
O'Hare, M.J., Soares, F.

TITLE Simpson, A. J.
Shotgun sequencing of
sequence tags

JOURNAL
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE 20202663

PUBMED 10737800

COMMENT	CONTACT: SIMPSON A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research
 Prof. Dr. J. Antonio Pariente 109 4 andar 01509-010 Sa

Rua Prof. Amilton Figueira 103, 4 andar, 01305-010, 00
Brazil

TEL: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human C

project. This entry can be seen in the following URL
<https://www.fda.gov/oc/foia/2017-2018-foia-requests>

(<http://www.ludwig.org.br/scripts/gethtml2.pl?cr=8422>)

200-070-a08&t3=2000-02-15&t4=1)
 000 primer. nuc 18 forward

seq primer: pac is forward
High quality sequence start: 28

High quality sequence stop: 496.

FEATURES	Location/Qualifiers
1. Location	1. Location
2. Qualifiers	2. Qualifiers
3. Other	3. Other
4. Other	4. Other
5. Other	5. Other
6. Other	6. Other
7. Other	7. Other
8. Other	8. Other
9. Other	9. Other
10. Other	10. Other
11. Other	11. Other
12. Other	12. Other
13. Other	13. Other
14. Other	14. Other
15. Other	15. Other
16. Other	16. Other
17. Other	17. Other
18. Other	18. Other
19. Other	19. Other
20. Other	20. Other
21. Other	21. Other
22. Other	22. Other
23. Other	23. Other
24. Other	24. Other
25. Other	25. Other
26. Other	26. Other
27. Other	27. Other
28. Other	28. Other
29. Other	29. Other
30. Other	30. Other
31. Other	31. Other
32. Other	32. Other
33. Other	33. Other
34. Other	34. Other
35. Other	35. Other
36. Other	36. Other
37. Other	37. Other
38. Other	38. Other
39. Other	39. Other
40. Other	40. Other
41. Other	41. Other
42. Other	42. Other
43. Other	43. Other
44. Other	44. Other
45. Other	45. Other
46. Other	46. Other
47. Other	47. Other
48. Other	48. Other
49. Other	49. Other
50. Other	50. Other
51. Other	51. Other
52. Other	52. Other
53. Other	53. Other
54. Other	54. Other
55. Other	55. Other
56. Other	56. Other
57. Other	57. Other
58. Other	58. Other
59. Other	59. Other
60. Other	60. Other
61. Other	61. Other
62. Other	62. Other
63. Other	63. Other
64. Other	64. Other
65. Other	65. Other
66. Other	66. Other
67. Other	67. Other
68. Other	68. Other
69. Other	69. Other
70. Other	70. Other
71. Other	71. Other
72. Other	72. Other
73. Other	73. Other
74. Other	74. Other
75. Other	75. Other
76. Other	76. Other
77. Other	77. Other
78. Other	78. Other
79. Other	79. Other
80. Other	80. Other
81. Other	81. Other
82. Other	82. Other
83. Other	83. Other
84. Other	84. Other
85. Other	85. Other
86. Other	86. Other
87. Other	87. Other
88. Other	88. Other
89. Other	89. Other
90. Other	90. Other
91. Other	91. Other
92. Other	92. Other
93. Other	93. Other
94. Other	94. Other
95. Other	95. Other
96. Other	96. Other
97. Other	97. Other
98. Other	98. Other
99. Other	99. Other
100. Other	100. Other

source	1. .498
--------	---------

```

.
.
/organism="Homo sapiens"

```

```

/mol_type="mRNA"
/mol_type="tRNA"
/mol_type="rRNA"

```

```

/db_xref="taxon:9606"
/db_xref="taxon:9606"

```

```
/dev_stage="Adults"
/clone lib="IT0036"
```

```

/crone_uid= p10030
/note="Organ: leiomyos: Vector: puc18: Site 1

```

Site 2: SmaI; A mini-library was made by cloning

derived from ORESTES PCR (U.S. Letters Patent

No. 196,716 - Ludwig Institute for Cancer Research

profiles into the pUC 18 vector. Reverse tra

tissue mRNA and cDNA amplification were per-

low stringency conditions."

ORIGIN

Alignment Scores:

```

Pred. No.: 0.115 Length: 498
Score: 11.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 25.58% Indels: 0
DB: 2 Gaps: 0

US-10-624-218-4 (1-43) x AW836776 (1-498)

QY 12 AsnThrAlaLeuArgThrGluValAlaGluLeu 22
Db 297 AACACAGCCCTGCGGACGGAGGTGGCGAGCTA 265

RESULT 11
BF660216
LOCUS 528 bp mRNA linear EST 20-DEC-2000
DEFINITION maa29a04.y1 NCI CGAP L110 Mus musculus cDNA clone IMAGE:3812262 5'
Similar to SW:TEF_RAT P41224 THYROTROPH EMBRYONIC FACTOR. [1] ;
mRNA sequence.
ACCESSION BF660216
VERSION BF660216.1 GI:11925350
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 528)
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
JOURNAL National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
COMMENT Tumor Gene Index
Unpublished (1997)
Other ESTs: maa29a04.x1
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-@email.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov

MGI:1454374
Seq primer: -40RP from Gibco
High quality sequence stop: 435.

FEATURES
source
1..528
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IWAGB:3812262"
/sex="female"
/dev_stage="10 weeks"
/lab_host="DH10B (TI phage-resistant)"
/clone_lib="NCI_CGAP L110"
/note="Organ: liver; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.6 kb. Library constructed by Life
Technologies."

ORIGIN
Alignment Scores:
Pred. No.: 0.121 Length: 528
Score: 11.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 25.58% Indels: 0
DB: 2 Gaps: 0

US-10-624-218-4 (1-43) x BF660216 (1-528)

QY 12 AsnThrAlaLeuArgThrGluValAlaGluLeu 22
Db 11925350

RESULT 12
BF660216
LOCUS 528 bp mRNA linear EST 20-DEC-2000
DEFINITION maa29a04.y1 NCI CGAP L110 Mus musculus cDNA clone IMAGE:3812262 5'
Similar to SW:TEF_RAT P41224 THYROTROPH EMBRYONIC FACTOR. [1] ;
mRNA sequence.
ACCESSION BF660216
VERSION BF660216.1 GI:11925350
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 528)
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
JOURNAL National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
COMMENT Tumor Gene Index
Unpublished (1997)
Other ESTs: maa29a04.x1
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-@email.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov

MGI:1454374
Seq primer: -40RP from Gibco
High quality sequence stop: 435.

FEATURES
source
1..528
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IWAGB:3812262"
/sex="female"
/dev_stage="10 weeks"
/lab_host="DH10B (TI phage-resistant)"
/clone_lib="NCI_CGAP L110"
/note="Organ: liver; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.6 kb. Library constructed by Life
Technologies."

ORIGIN
Alignment Scores:
Pred. No.: 0.121 Length: 528
Score: 11.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 25.58% Indels: 0
DB: 2 Gaps: 0

US-10-624-218-4 (1-43) x BF660216 (1-528)

QY 12 AsnThrAlaLeuArgThrGluValAlaGluLeu 22
Db 11925350

RESULT 13
AW836791/c
LOCUS 555 bp mRNA linear EST 18-MAY-2000
DEFINITION QV1-LT0036-150200-070-f06 LT0036 Homo sapiens cDNA, mRNA sequence.
ACCESSION AW836791
VERSION AW836791.1 GI:7930765
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 555)
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 555)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,

```

Goldman, G.H., Carvalho, A.F., Matekuma, A., Baia, G.S., Simpson, D.H., Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.

Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000).

20202663

10737800

COMMENT

Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL (<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=4t2-QV1-LT0036-150>)

200-070-f06&t3=2000-02-15&t4=1

Seq primer: puc 18 forward

High quality sequence stop: 555.

FEATURES

source

1..555

Location/Qualifiers

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/dev_stage="Adult"

/clone_lib="LT0036"

/note="Organ: leiomyos; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN

Alignment Scores:
Pred. No.: 0.127 Length: 555
Score: 11.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 25.58% Indels: 0
DB: 2 Gaps: 0

US-10-624-218-4 (1-43) x AW836791 (1-555)

QY 12 AsnThrAlaLeuArgThrGluValAlaGluLeu 22

Db 344 AACACAGCCCTCGGACGGAGTGCGCGAGCTA 312

RESULT 14

BU311019

LOCUS

DEFINITION

BU311019

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

12445392

sequence.

sequence.

sequence.

sequence.

sequence.

sequence.

sequence.

sequence.

sequence.

COMMENT

Contact: Simon Hubbard
Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology
(UMIST)

PO Box 88, Manchester, M60 1QD, UK

Tel: 01612008930

Fax: 01612360409

Email: Simon.Hubbard@umist.ac.uk.

Location/Qualifiers

1..557

source

/organism="Gallus gallus"

/mol_type="mRNA"

/strain="Compton line 151"

/db_xref="taxon:9031"

/clone="ChEST509n17"

/sex="Female"

/dev_stage="adult"

/lab_host="DH10B"

/clone_lib="CSEQCCHN61"

/note="Organ: heart; Vector: pBluescript II KS(+); Site_1: EcoRI; Site_2: NotI; This normalized library was constructed from 1 million independent clones. cDNA synthesis was initiated using an oligo(dT) primer, using methylated C in the first strand synthesis reaction. Following this first strand reaction, double-stranded cDNA was blunt-ended, ligated to NotI adapters, digested with EcoRI, size-selected, and cloned into the NotI and EcoRI compatible sites of a custom modified MCS of the pBluescript (KS+) vector. The library was normalized in 2 rounds using conditions adapted from Soares et al., PNAS (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6 (1996): 791, except that a significantly longer reannealing hybridization was used."

ORIGIN

Alignment Scores:
Pred. No.: 0.127 Length: 557
Score: 11.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 25.58% Indels: 0
DB: 5 Gaps: 0

US-10-624-218-4 (1-43) x BU311019 (1-557)

QY 12 AsnThrAlaLeuArgThrGluValAlaGluLeu 22

Db 141 AATACGCCCTCGGACGGAGTGTCAGAGCTG 173

RESULT 15

CO794673

LOCUS

DEFINITION

CO794673

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1 (bases 1 to 563)

1 (bases 1 to 563)

1 (bases 1 to 563)

1 (bases 1 to 563)

1 (bases 1 to 563)

1 (bases 1 to 563)

1 (bases 1 to 563)

1 (bases 1 to 563)

1 (bases 1 to 563)

1 (bases 1 to 563)

1 (bases 1 to 563)

1 (bases 1 to 563)

1 (bases 1 to 563)

1 (bases 1 to 563)

1 (bases 1 to 563)

1 (bases 1 to 563)

CO794673 563 bp mRNA linear EST 05-AUG-2004
NT144B_B07 St18-22 Neural tube (NT) Ambystoma mexicanum cDNA 5', similar to vitellogenin binding protein (VBP), beta/beta isoform (Gallus, mRNA sequence).

CO794673

CO794673.1 GI:51010644

EST

AMBYSTOMA MEXICANUM (AXOLOTL)

AMBYSTOMA MEXICANUM

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Amphibia; Batrachia; Caudata; Salamandroidea; Ambystomidae;

Ambystoma.

1 (bases 1 to 563)

1 (bases 1 to 563)

1 (bases 1 to 563)

1 (bases 1 to 563)

1 (bases 1 to 563)

1 (bases 1 to 563)

1 (bases 1 to 563)

1 (bases 1 to 563)

1 (bases 1 to 563)

1 (bases 1 to 563)

1 (bases 1 to 563)

1 (bases 1 to 563)

1 (bases 1 to 563)

1 (bases 1 to 563)

1 (bases 1 to 563)

1 (bases 1 to 563)

1 (bases 1 to 563)

Habermann, B., Bebin, A.G., Herklotz, S., Volkmer, M., Eckelt, K., Fehlike, K., Epperlein, H.H., Schackert, H.K., Wiebe, G. and Tanaka, E.M. An Ambystoma mexicanum EST sequencing project: Analysis of 17,352 expressed sequence tags from embryonic and regenerating blastema cDNA libraries
Genome Biol. (2004) In press
Contact: Elly M. Tanaka
Tanaka Lab
Max Planck Institute of Molecular Cell Biology and Genetics,

Dresden
Pfortenhauerstrasse 108,01307 Dresden, Germany
Tel: 0049 351 210 2620
Fax: 0049 351 210 1489
Email: tanaka@mpi-cbg.de
Plate: NT144B row: 07 column: B
Seq primer: GCA CAT TAG GCC TAT TTA GGT GAC A.

FEATURES
source
1..563
/organism="Ambystoma mexicanum"
/mol_type="mRNA"
/db_xref="taxon:8296"
/tissue_type="Neural Tube, Notochord, Somites"
/cell_type="Includes Neural tube, notochord, somites"
/dev_stage="Stage 18-22"
/clone_lib="St18-22 Neural tube (NT)"
/notes="vector: pCMVSPORT6; Site_1: NotI; Site_2: SalI;
Unnormalized cDNA plasmid library prepared by Invitrogen.
Size fractionated mRNA was polydt primed and cloned into
NotI-SalI site of pCMVSPORT6. Bacterial host is
EMDH10B-TONA. Average insert size is 1.5 kb.
TAG_Lib=NT"

ORIGIN

Alignment Scores:
Pred. No.: 0.129 Length: 563
Score: 11.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 25.58% Indels: 0
DB: 7 Gaps: 0

US-10-624-218-4 (1-43) x C0794673 (1-563)

Qy 12 AsnThrAlaLeuArgThrGluValAlaGluLeu 22
Db 95 AATACAGCCTTGGCACAGAGTGGCTGAGCTG 127

Search completed: July 27, 2005, 16:41:26
Job time : 2993.69 secs